

# Investigating Affective Responses to Video Game Events: An exploratory study utilising psychophysiology

Benjamin James Geelan, BIS (Hons)

Submitted in fulfilment of the requirements for the degree of  
**Doctor of Philosophy**



UNIVERSITY  
OF TASMANIA

**University of Tasmania**

November 2017

### **Declaration of Originality**

This thesis contains no material which has been accepted for a degree or diploma by the University or any other institution, except by way of background information and duly acknowledged in the thesis, and to the best of the candidate's knowledge and belief no material previously published or written by another person except where due acknowledgement is made in the text of the thesis, nor does the thesis contain any material that infringes copyright

.....

Benjamin James Geelan

9<sup>th</sup> November 2017

### **Authority of Access**

This thesis may be made available for loan and limited copying in accordance with the *Copyright Act 1968*.

.....

Benjamin James Geelan

9<sup>th</sup> November 2017

### **Statement of Ethical Conduct**

The research associated with this thesis abides by the international and Australian codes on human and animal experimentation, the guidelines by the Australian Government's Office of the Gene Technology Regulator and the rulings of the Safety, Ethics and Institutional Biosafety Committees of the University.

.....

Benjamin James Geelan

9<sup>th</sup> November 2017

### Abstract

This thesis presents the results of three exploratory studies aimed at developing a psychophysiological method capable of examining the influence that events occurring during video game play had upon the affective responses of players, including the examination of measurement effects associated with the use of psychophysiological measures.

The recent explosion of video games as a form of entertainment has consequentially led to the rapid expansion of research investigating video games and phenomena associated. However, many aspects of video games remain under examined or poorly understood, not the least of which is the relationship between video games and emotions. The study of emotions and behaviours within video games has been traditionally dominated by subjective and qualitative measures, which are valuable for developing understandings about experiences and behaviours. However, the nature of these methods and their use mean that they are generally unable to provide reliable or meaningful insights into events or stimuli that occur during normal, uninterrupted, dynamic (Klasen *et al.*, 2008; Mandryk and Atkins, 2007).

The relatively recent and rapidly growing interdisciplinary science that attempts to avoid these issues is psychophysiology, which involves objectively measuring quantitative aspects of the human body to assess those that relate to the experience of emotions and behaviour (Cacioppo *et al.*, 2007; Cacioppo and Tassinary, 1990). This allows studies to collect objective data on emotions covertly, and without the biases present in subjective measures (Ijsselstein, de Kort, and Poels, 2008; Drachen *et al.*, 2010).

However, despite the recent successes of psychophysiology and the promise it holds for the investigation of event-related phenomena occurring during video game play owing to its temporal precision and continuous measurements, relatively few studies attempt to examine real-time game events (Kivikangas *et al.*, 2011; Cowley *et al.*, 2016). Additionally, while the use of psychophysiological measures may be relatively

covert, the placement of electrodes upon the body and face of participants may fundamentally alter the experience of video game play. However, changes in player experience stemming from the use of psychophysiology measures do not appear to have been sufficiently examined within the video game domain.

The research presented within this thesis therefore aims to address these gaps in knowledge, and adapts a high-resolution objective psychophysiological method based upon previous studies conducted by Mandryk and Atkins (2007), Mandryk *et al.* (2008), Nacke and Lindley (2009), Nacke *et al.* (2009), Ravaja and Kivikangas (2008), Ravaja *et al.* (2006), and with reference to guidelines presented by Ravaja *et al.* (2009;2011) and Cowley *et al.* (2016). Through this method, this research project identifies and presents three areas where novel contributions to the domain of video game research have been made:

- **A novel adaptation of existing psychophysiological methods** to measure and analyse event-related affective responses to observed video game events. Ten events were identified during video game play and analysed against five modelled emotions, with differences from the mean affective experience examined for each. These revealed insights into the experiences and game mechanics underlying these affective responses, as well as the events that elicited them. This method therefore allows researchers and developers to measure and assess responses to events occurring within their normal environment, without requiring that they be examined in isolation from normal game experiences. This facilitates improvements to both the psychological understandings of game stimuli, as well as allowing developers to better assess the influence of design decisions.
- **A comparison between events** that occurred for both the player and the opponent, and an examination of differences in affective responses elicited in real-time and on a per-event basis. This provided new insights into the interplay between the personal experience of events and events occurring to

other players. This method therefore allows researchers and developers to reliably measure and assess the influence of game mechanics or event interactions in real time, including differences produced through the presence of social interactions.

- **An examination of measurement effects** associated with the use of psychophysiological equipment, revealing that the presence of equipment significantly reduced self-reported immersion, and significantly increased self-reported challenge, as well as perceptions of both positive and negative emotions. This provided novel insights into measurement effects elicited through use of psychophysiological equipment, with indications that participants became distracted and introspective when electrodes were placed on their bodies. This method therefore provides preliminary insights into the changes in experience that physiological measures may produce, which is valuable for researchers and developers considering the use of physiological measures for the purposes of measuring or assessing player emotions and experience.

## Acknowledgements

*"I blame all of you. Writing this book has been an exercise in sustained suffering. The casual reader may, perhaps, exempt herself from excessive guilt, but for those of you who have played the larger role in prolonging my agonies with your encouragement and support, well...you know who you are, and you owe me."*

-- Brendan Pietsch --

First of all, I would like to express my gratitude to my supervisors, Dr. Kristy de Salas and Dr. Ian Lewis, who provided inspiration to motivate my work, encouragement when I needed it most, guidance even when I did not believe I needed it, and support through the worst of times.

I would like to thank my research colleagues whom I had the fortune of sharing my lab with: Anthony Smith, Aran Cauchi-Saunders, Adam Zulkifly, and Lindsay Wells. There is no doubt that your feedback, deplorable excessive of personality, and the hours of Duck Game (and Party Golf) were instrumental in helping me through to completion.

To all my family and friends who I have unfortunately neglected in no small, you have my deepest gratitude for your interminable patience, (mostly) understanding, and support. I would like to thank you all for lighting a path through your friendship, laughter, guidance, and happiness throughout the dark years.

It takes a special kind of crazy for a someone outside of supervisors to demand to see drafts, and to offer honest advice and feedback. Therefore, I would like to thank Sharon Steinemann for her tireless and selfless support, for her inspiration and encouragement, for holding me accountable, and for always understanding.



## Table of Contents

<b>Declaration .....</b>	<b>ii</b>
<b>Abstract.....</b>	<b>v</b>
<b>Acknowledgements .....</b>	<b>viii</b>
<b>Table of Figures .....</b>	<b>xv</b>
<b>List of Tables.....</b>	<b>xxiii</b>
<b>Chapter 1: Introduction.....</b>	<b>1</b>
1.1. Project Background .....	1
1.2. Purpose of this Research Project .....	2
1.3. Research Questions and Strategy .....	5
1.3.1. Research Questions .....	5
1.3.2. Research Methods for this Thesis.....	6
1.4. Context of the Research.....	6
1.5. Summary of Contributions .....	7
1.6. Layout of Chapters .....	8
1.6.1. Chapter 1: Introduction .....	8
1.6.2. Chapter 2: Literature Review .....	8
1.6.3. Chapter 3: Methodology.....	8
1.6.4. Chapter 4: Results .....	9
1.6.5. Chapter 5: Discussions .....	9
1.6.6. Chapter 6: Research Summary .....	9
1.7. Chapter Reflections .....	9
<b>Chapter 2: Literature Review.....</b>	<b>10</b>
2.1. Chapter Introduction.....	10
2.2. Domain 1: Video Game Research.....	10
2.2.1. Social Video Game Play Experiences .....	12

2.2.2. Existing Measurement of Game Experiences .....	14
2.2.3. Shift towards Psychophysiology .....	16
2.3. Domain 2: Psychophysiology .....	17
2.3.1. Overview .....	17
2.3.2. Psychophysiological Signals .....	17
2.3.3. Arousal .....	39
2.3.4. Affective States .....	41
2.3.5. Summary of Psychophysiology .....	44
2.4. Junction of Domains: Video Game Research Using Psychophysiology .....	46
2.4.1. Event-Related Responses .....	47
2.4.2. Measurement Effects of Psychophysiology on Game Experiences .....	49
2.5. Opportunity statement .....	51
<b>Chapter 3: Methodology .....</b>	<b>54</b>
3.1. Chapter Introduction .....	54
3.2. Introduction .....	54
3.2.1. Research Paradigm .....	55
3.2.2. Research Objectives .....	61
3.2.3. Research Questions .....	62
3.2.4. Data Requirements .....	65
3.2.5. Ethics .....	68
3.3. Instruments .....	69
3.3.1. Psychophysiological devices .....	69
3.3.2. The Game .....	80
3.3.3. Video Recording .....	84
3.3.4. Baselines .....	85

3.4.	Primary Measures: Physiological .....	86
3.4.1.	Brain Activity: Electroencephalography (EEG).....	86
3.4.2.	Affective Valence: Facial Electromyography (EMG) .....	87
3.4.3.	Physiological Arousal: Heart Rate Photoplethysmography (PPG) .....	90
3.4.4.	Physiological Arousal: Galvanic Skin Response (GSR).....	92
3.4.5.	Primary Measures: Summary.....	93
3.5.	Secondary Measures: Observational and Subjective Data .....	94
3.5.1.	Subjective Measures of Player Experiences: Game Experience Questionnaire (GEQ) .....	94
3.5.2.	Subjective Measures of Affective Responses: .....	96
3.5.3.	Demographic Survey .....	98
3.5.4.	Secondary Measures Summary .....	98
3.6.	Methodology: Pilot Studies .....	99
3.6.1.	Method and Research Design .....	99
3.6.2.	Participant Recruitment.....	99
3.6.3.	Pilot 1: Basic physiological arousal through Heartrate.....	100
3.6.4.	Pilot 2: EEG brain activity .....	106
3.6.5.	Pilot 3: Trial of primary procedure.....	122
3.6.6.	Methodology: Pilot Study Summary .....	123
3.7.	Methodology: Main Study.....	123
3.7.1.	Method and Research Design .....	124
3.7.2.	Participant Recruitment.....	125
3.7.3.	Experimental Setup and Procedures .....	126
3.7.4.	Data Processing.....	130
3.7.5.	Methodology: Primary Study Summary.....	140
3.8.	Modelling of Emotions from Physiological Sources .....	140

3.8.1. Fuzzy Logic Approaches .....	142
3.8.2. The Mandryk Model of Emotion.....	143
3.8.3. Summary of Fuzzy Model of Emotion.....	149
3.9. Catalogue of Events.....	150
3.9.1. Overview .....	150
3.9.2. Novel Design: Catalogue of Events (CoE).....	152
3.9.3. Registering and Coding Events into the Catalogue.....	154
3.9.4. Event-Related Windows for Physiological Data.....	155
3.9.5. Limitations.....	155
3.9.5. Summary .....	157
3.10 . Methodology Summary.....	158
<b>Chapter 4: Results.....</b>	<b>158</b>
4.1. Layout of Chapter / Structure .....	158
4.2. Research Question 1 .....	160
4.2.1. Overall Objective Game Play Experience.....	162
4.2.2. Comparison of Events with Overall Game Experience .....	167
4.2.3. Summary of Results for Research Question 1 .....	207
4.3. Research Question 2 .....	209
4.3.1. Subjective Game Experience: Equipment On vs. Equipment Off .....	211
4.3.2. Comparing Objective and Subjective Measures .....	225
4.3.3. Summary of Results for Research Question 2 .....	230
4.4. Summary of Results Chapter.....	231
<b>Chapter 5: Discussions .....</b>	<b>232</b>
5.1. Layout of Chapter / Structure .....	232
5.2. Introduction.....	232

5.3.	Research Question 1 .....	233
5.3.1.	Mean Affective Game Experiences .....	236
5.3.2.	Event-Related Affective Responses .....	242
5.3.3.	Summary of Research Question 1 .....	262
5.4.	Research Question 2 .....	263
5.4.1.	Relationship between Subjective and Objective Measures .....	264
5.4.2.	Measurement Effects.....	267
5.4.3.	Summary of Research Question 2 .....	274
5.5.	Summary of Discussion Chapter .....	275
<b>Chapter 6:</b>	<b>Research Summary .....</b>	<b>277</b>
6.1.	Introduction.....	277
6.2.	Summary of the Research .....	277
6.2.1.	Implications of the Research Findings .....	280
6.3.	Possibilities for Future Work .....	285
6.4.	Concluding Reflections .....	288
<b>References.....</b>		<b>289</b>
<b>Appendices.....</b>		<b>320</b>
Appendix A:	Alternative Video Games Considered.....	320
1:	Super Monkey Ball 2 .....	320
2:	Left 4 Dead 2 .....	321
3:	Mario Party 10 .....	322
Appendix B:	Alternative Subjective Measures of Player Experience .....	324
1:	Player Experience of Needs Satisfaction (PENS).....	324
2:	Social Presence in Gaming Questionnaire (SPGQ).....	325
Appendix C:	Alternative Subjective Measures of Affect .....	326

1: Profile of Mood States (POMS) .....	326
Appendix D: Fuzzy Logic Rules Used .....	328
1. Mapping Physiological Signals to AV Space .....	328
2. Mapping AV Space Data to 5 Classifications of Emotions.....	329
Appendix E: The Process of Registering and Coding Events within the CoE.....	332
1. Video Recordings: Facial Expressions.....	332
2. Video Recordings: Social Interactions .....	334
3. Video Recordings: Game Events .....	335
4. Model of Emotions: Affective Responses .....	336

## Table of Figures

<i>Figure 1: The use of traditionally subjective and qualitative instruments within studies of video games, with a notable gap in instruments that are both objective and quantitative (Nacke, 2009).</i>	16
<i>Figure 2: Functional brain regions and their associated mental activities and processes (Sukel, 2011).</i>	19
<i>Figure 3: Raw EEG signals for multiple channels and frequencies, as recorded using the BioRadio device and software (GLNeuroTechnologies, 2014).</i>	21
<i>Figure 4: EEG heat-maps showing brain region activation and deactivation, created through processing the raw EEG signal data (Koelstra and Muhl, 2011).</i>	22
<i>Figure 5: Heart rate variations in response to video game stimuli, showing peaks and dips in the heart rate as different components within the game are experienced. Graph is an excerpt of data collected from one participant during pre-pilot stage of this research project.</i>	24
<i>Figure 6: Diagram showing correct ECG electrode placement on the chest for a 7-lead ECG system (STAT Medical Education, 2014).</i>	26
<i>Figure 7: Raw ECG signals and the HR and RR extrapolated from the ECG signal, as recorded through a BioRadio device and software (GLNeuroTechnologies, 2014).</i>	27
<i>Figure 8: A portable Pulse Oximeter, illustrating the finger clip and a wrist-device for displaying the heart rate and level of blood.</i>	28
<i>Figure 9: Electrodermal responses to video game stimuli, shown over a 60 second period after the experience of a stimuli (Galvanic Skin Response, n.d.).</i>	30
<i>Figure 10: A GSR device attached to the fingers of a participant (YCAN Australia, 2014).</i>	32
<i>Figure 11: Different facial expressions, and their associated emotional states (OMRON, 2014).</i>	34
<i>Figure 12: EMG electrodes being used on a participant, showing the placement on the three most commonly measured facial muscles (brow, eyelid, and cheek) (Texas Tech, 2014).</i>	37

<i>Figure 13: Russell's two-dimensional space with Arousal and Valence, and the emotions associated with different coordinates within the model (Ambinder, 2011).</i>	43
<i>Figure 14: BioRadio device on arm (GLNeuroTechnologies, 2014).</i>	70
<i>Figure 15: Comparison of 64 channel laboratory EEG device against a 16 channel portable EPOC+ EEG device (Duvina et al., 2013).</i>	71
<i>Figure 16: EPOC+ EEG headset placed on head of participant (Emotiv, 2015).</i>	72
<i>Figure 17: Traditional EEG and EMG electrodes placed on head of participant.</i>	72
<i>Figure 18: EEGLAB visualization of time/frequency event-related EEG data (EEGLAB, 2015).</i>	75
<i>Figure 19: LEDALAB visualization of electrodermal activity over time during event stimuli (LEDALAB, 2015).</i>	76
<i>Figure 20: EMGLAB visualization of electromyographic activity over time (EMGLAB, 2015).</i>	77
<i>Figure 21: HRVAS visualisation of heart rate activity over time, including breakdowns of instantaneous HR and variability across a number of measures (Ramshur, 2010).</i>	79
<i>Figure 22: One of the competitive game modes in Tetris Party Deluxe, facilitating direct competition between players (Tetris, 2010).</i>	82
<i>Figure 23: One of the cooperative game modes in Tetris Party Deluxe, facilitating direct cooperation between players (Tetris, 2010).</i>	83
<i>Figure 24: Illustration of facial muscles, with Corrugator supercilii highlighted (Janis et al., 2007).</i>	88
<i>Figure 25: Illustration of facial muscles, with Zygomaticus major highlighted (Janis et al., 2007).</i>	89
<i>Figure 26: Example of ECG electrode placement on participant.</i>	90
<i>Figure 27: Pulse Oximeter photoplethysmography (PPG) placement on participant.</i>	91
<i>Figure 28: GSR electrode placement on participant.</i>	93
<i>Figure 29: The Self-Assessment Manikin (SAM) from the IAPS norms (Lang, Greenwald, and Bradley, 1989). The rows (from top to bottom) require individuals to rate their perceived pleasure, arousal, and dominance.</i>	97
<i>Figure 30: Boxplot used to assess for outliers, with none evidenced.</i>	103



<i>Figure 31: Visualisation of Event Related EEG data in Brainstorm, illustrating region of brain activation and functional connectivity. ....</i>	<i>109</i>
<i>Figure 32: Boxplot used to assess for outliers, with none evidenced. ....</i>	<i>111</i>
<i>Figure 33: Mean differences in Theta activity between Competitive Play, Cooperative Play, and Solo Play. ....</i>	<i>113</i>
<i>Figure 34: Mean differences in Alpha activity between Competitive Play, Cooperative Play, and Solo Play. ....</i>	<i>115</i>
<i>Figure 35: Mean differences in Beta activity between Competitive Play, Cooperative Play, and Solo Play. ....</i>	<i>117</i>
<i>Figure 36: Mean differences in Gamma activity between Competitive Play, Cooperative Play, and Solo Play.....</i>	<i>119</i>
<i>Figure 37: Diagram of experimental room setup.....</i>	<i>127</i>
<i>Figure 38: Facial EMG muscle locations and associated electrode placements (Boxtel, 2010). ....</i>	<i>134</i>
<i>Figure 39: The Self-Assessment Manikin (SAM) from the IAPS norms (Lang, Greenwald, and Bradley, 1989). The rows (from top to bottom) require individuals to rate their perceived pleasure, arousal, and dominance. ....</i>	<i>136</i>
<i>Figure 40: Formula used to normalise SAM results into percentages of maximum. .</i>	<i>137</i>
<i>Figure 41: Formula used to normalise GEQ results into percentages of maximum. .</i>	<i>139</i>
<i>Figure 42: The modified Affect Grid used by Mandryk et al. (2006) to model physiological signals and interpret arousal and valence. ....</i>	<i>143</i>
<i>Figure 43: The mamdani fuzzy system used by Mandryk et al. (2006) to classify four physiological signals into arousal and valence. ....</i>	<i>145</i>
<i>Figure 44: 3D models depicting how the 22 membership rules map physiological signals into varying degrees of arousal and valence. From Mandryk et al. (2006)...</i>	<i>146</i>
<i>Figure 45: The Emotion Modelling mamdani fuzzy system used by Mandryk et al. (2006) to classify arousal and valence (AV space) into five emotional outputs. Each input and output contains a schematic representing the location and form of each respective membership function.....</i>	<i>147</i>

<i>Figure 46: Representation of modelled emotions on the Affect Grid, which classifies emotions based upon affect and valence within the AV space previously discussed. From Mandryk et al. (2006). .....</i>	<i>148</i>
<i>Figure 47: Box plot illustrating the distribution of aggregate data for arousal and valence. Note that no outliers were identified. ....</i>	<i>164</i>
<i>Figure 48: Affect Grid modelling the mean arousal and valence scores, illustrating that the overall game experience was slightly arousing with a positive valence.....</i>	<i>165</i>
<i>Figure 49: Box plot illustrating the distribution of aggregate data for each modelled emotion. Note that no outliers were identified. ....</i>	<i>166</i>
<i>Figure 50: Box plot illustrating the distribution of AV space and emotions occurring during Speed Increase events. Note that no outliers are identified. ....</i>	<i>169</i>
<i>Figure 51: Affect Grid modelling the mean arousal and valence scores occurring during Speed Increase events, illustrating that the affective responses were moderately arousing.....</i>	<i>170</i>
<i>Figure 52: Box plot illustrating the distribution of AV space and emotions occurring during Non-Useful Tetriminos Received events. Note that no outliers are identified. ....</i>	<i>173</i>
<i>Figure 53: Affect Grid modelling the mean arousal and valence scores occurring during Non-Useful Tetriminos Received events, illustrating that the affective responses were minimal but slightly negatively valenced.....</i>	<i>174</i>
<i>Figure 54: Box plot illustrating the distribution of AV space and emotions occurring during Correct Tetriminos Placement events. Note that no outliers are identified...177</i>	<i>177</i>
<i>Figure 55: Affect Grid modelling the mean arousal and valence scores occurring during Correct Tetriminos Placement events, illustrating that the affective responses were slightly moderately arousing and positively valenced. ....</i>	<i>178</i>
<i>Figure 56: Box plot illustrating the distribution of AV space and emotions occurring during Incorrect Tetriminos Placement events. Note that no outliers were identified. ....</i>	<i>181</i>
<i>Figure 57: Affect Grid modelling the mean arousal and valence scores occurring during Incorrect Tetriminos Placement events, illustrating that the affective responses were minimal but slightly negatively valenced.....</i>	<i>182</i>

<i>Figure 58: Box plot illustrating the distribution of AV space and emotions occurring during Player Clears Multiple Lines (“Tetris”) events. Note that no outliers are identified. ....</i>	<i>185</i>
<i>Figure 59: Affect Grid modelling the mean arousal and valence scores occurring during Multiple Line Clear (“Tetris”) events, illustrating that the affective responses were minimal but slightly negatively valenced.....</i>	<i>186</i>
<i>Figure 60: Box plot illustrating the distribution of AV space and emotions occurring during Player Win events. Note that no outliers are identified. ....</i>	<i>189</i>
<i>Figure 61: Affect Grid modelling the mean arousal and valence scores occurring during Player Win events, illustrating that the affective responses were highly arousing and positively valenced. ....</i>	<i>190</i>
<i>Figure 62: Box plot illustrating the distribution of AV space and emotions occurring during Player Loss events. Note that no outliers are identified.....</i>	<i>193</i>
<i>Figure 63: Affect Grid modelling the mean arousal and valence scores occurring during Player Loss events, illustrating that the affective responses were arousing and very low positively valenced.....</i>	<i>194</i>
<i>Figure 64: Box plot illustrating the distribution of AV space and emotions occurring during Opponent Correct Tetriminos Placement events. Note that no outliers are identified. ....</i>	<i>197</i>
<i>Figure 65: Affect Grid modelling the mean arousal and valence scores occurring during Opponent Correct Tetriminos Placement events, illustrating that the affective responses were minimal but slightly positively valenced. ....</i>	<i>198</i>
<i>Figure 66: Box plot illustrating the distribution of AV space and emotions occurring during Opponent Incorrect Tetriminos Placement events. Note that no outliers are identified. ....</i>	<i>201</i>
<i>Figure 67: Affect Grid modelling the mean arousal and valence scores occurring during Opponent Incorrect Tetriminos Placement events, illustrating that the affective responses were moderate and slightly positively valenced. ....</i>	<i>202</i>
<i>Figure 68: Box plot illustrating the distribution of AV space and emotions occurring during Opponent Multiple Line Clear (“Tetris”) events. Note that no outliers are identified. ....</i>	<i>205</i>

<i>Figure 69: Affect Grid modelling the mean arousal and valence scores occurring during Opponent Multiple Line Clear (“Tetris”) events, illustrating that the affective responses were minimal but slightly negatively valenced.....</i>	<i>206</i>
<i>Figure 70: Overview of the modelled AV space for each of the game events analysed within the results section of this project. In-depth analysis, including effect size, strength, and significance of differences is presented in the previous sections. ....</i>	<i>208</i>
<i>Figure 71: Overview of the modelled Emotional categories for each of the game events analysed within the results section of this project. In-depth analysis, including effect size, strength, and significance of differences is presented in the previous sections. ....</i>	<i>209</i>
<i>Figure 72: Box plot illustrating the distribution of subjective Game Experience Questionnaire CORE components. Note that no outliers are identified, including those for data not presented. ....</i>	<i>212</i>
<i>Figure 73: Chart illustrating mean differences between Equipment On and Equipment Off conditions for each of the GEQ CORE components.....</i>	<i>214</i>
<i>Figure 74: Chart illustrating mean differences between Equipment On and Equipment Off conditions for each of the GEQ SOCIAL components. ....</i>	<i>217</i>
<i>Figure 75: Box plot illustrating the distribution of subjective SAM components for both Equipment On and Equipment Off conditions. Note that no outliers are identified. ....</i>	<i>220</i>
<i>Figures 76 and 77: Affect Grids modelling the mean subjective arousal and valence scores for Equipment On (Left Affect Grid) and Equipment Off (Right Affect Grid). Valence stays similar for both conditions, however arousal appeared to increase when Equipment was not used. ....</i>	<i>221</i>
<i>Figure 78: Chart illustrating mean differences between Equipment On and Equipment Off conditions for each of the SAM components. ....</i>	<i>222</i>
<i>Figure 79: Chart illustrating mean cumulative differences between Equipment On and Equipment Off conditions for both Positive and Negative Expressions. ....</i>	<i>224</i>
<i>Figure 80: Chart illustrating mean differences in Arousal and Valence under Equipment On and Equipment Off conditions, for both Objective and Subjective measures. ....</i>	<i>227</i>

<i>Figure 81: Scatterplot modelling Subjective Arousal against Objective Arousal to investigate the relationship between measures. Note a general linear relationship.</i>	228
<i>Figure 82: Scatterplot modelling Subjective Valence against Objective Valence to investigate the relationship between measures. Note a general linear relationship.</i>	228
<i>Figure 83: Overview of the subjective measures analysed within the results section of this project for both Equipment On and Equipment Off conditions. All data in mean % of maximum score except for Observed Positive and Negative Expressions, which are in mean seconds of expression duration. In-depth analysis, including effect size, strength, and significance of differences is presented in the previous sections.</i>	231
<i>Figure 70 (repeated): Overview of the modelled AV space for each of the game events analysed within the results section of this project. In-depth analysis, including effect size, strength, and significance of differences is presented in the Results chapter.</i>	235
<i>Figure 71(repeated): Overview of the modelled Emotional categories for each of the game events analysed within the results section of this project. In-depth analysis, including effect size, strength, and significance of differences is presented in the Results chapter.</i>	236
<i>Figure 13 (repeated): A sample Circumplex model of affect representing the two-dimensional space of Arousal and Valence, and the emotions associated with different coordinates within the model (Ambinder, 2011).</i>	238
<i>Figure 87 (repeated): Classification of the five categories of emotions based on their respective arousal and valence. From Mandryk et al. (2007).</i>	243
<i>Figure 88: Affect Grid illustrating the mean Arousal and Valence for Game Speed Increase events.</i>	244
<i>Figure 89: Nacke and Lindlay's (2009) two-dimensional four-channel model of flow, which they adapted from Csíkszentmihályi (1975) and Ellis, Voelkl, and Morris (1994).</i>	245
<i>Figure 90: Affect Grid illustrating the mean Arousal and Valence for Non-Useful Tetriminos events.</i>	248

<i>Figure 91: Non-Useful Tetriminos Received events placed in order of occurrence for a single participant, illustrating changes in event-response over time. ....</i>	<i>250</i>
<i>Figure 92: Affect Grid modelling the mean arousal and valence scores occurring during Correct Tetriminos Placement events when completed by Players (Left) and Opponents (Right), illustrating that the affective responses were moderately arousing and positively valenced. ....</i>	<i>251</i>
<i>Figure 93: Affect Grid modelling the mean arousal and valence scores occurring during Incorrect Tetriminos Placement events when completed by Players (Left) and Opponents (Right). ....</i>	<i>254</i>
<i>Figure 94: Affect Grid modelling the mean arousal and valence scores occurring during Multiple Line Clear events as experienced by Players (Left) and Opponents of Players (Right). ....</i>	<i>257</i>
<i>Figure 95: Affect Grid modelling the mean arousal and valence scores occurring during Player Win (Left) and Player Loss (Right) events, illustrating that the affective responses were considerably different. ....</i>	<i>259</i>
<i>Figure 96: Chart illustrating mean differences in Arousal and Valence for both Objective and Subjective measures. ....</i>	<i>265</i>
<i>Figure 97: Affect Grids illustrating differences between the mean Arousal and Valence scores for both Objective Measures (Left grid) and Subjective Measures (Right grid). ....</i>	<i>267</i>
<i>Figure 98: Chart illustrating mean differences between Equipment On and Equipment Off conditions for each of the GEQ CORE components. ....</i>	<i>270</i>
<i>Figure 99: Chart illustrating mean differences between Equipment On and Equipment Off conditions for each of the GEQ SOCIAL components. ....</i>	<i>271</i>
<i>Figure 100: A racing level in Super Monkey Ball 2, showing a series of platforms and obstacles (Sega, 2003). ....</i>	<i>321</i>
<i>Figure 101: Four players working together to survive an attacking horde of infected in Left 4 Dead 2 (Valve, 2009). ....</i>	<i>322</i>
<i>Figure 102: Players competing in a mini-game in Mario Party 10 (Nintendo, 2015). ....</i>	<i>323</i>

*Figure 103: A sample Coding Sheet used within this project, depicting recordings of the time logs of each expression, the calculated duration, estimated intensity of expression, and areas to code expressions as either positively or negatively valenced. Adapted from Kring and Sloan (1991). ....333*

### List of Tables

Table 1: Summary of Research Questions and their respective data requirements, explained further below. ....	66
Table 2: Descriptive statistics for Pilot 1, showing mean data +- standard deviation. ....	102
Table 3: Shapiro-Wilk's Test of Normality, evidencing normal distribution.....	103
Table 4: Mauchly's Test of Sphericity, evidencing that sphericity has not been violated. ....	104
Table 5: Output of one-way within-subjects ANOVA, evidencing a significant difference in Heart Rate between Play conditions. ....	104
Table 6: Multiple paired-samples t-tests with a Bonferroni adjustment, evidencing statistically significant differences in means between HR_Solo and HR_Cooperation Play conditions. ....	105
Table 7: Descriptive statistics for Pilot 2, showing mean data +- standard deviation. ....	111
Table 8: Output of one-way within-subjects ANOVA for relative Theta frequency activity, evidencing a significant difference across the Play conditions.....	112
Table 9: Multiple paired-samples t-tests with a Bonferroni adjustment, evidencing statistically significant differences in means. ....	113
Table 10: Output of one-way within-subjects ANOVA for relative Alpha frequency activity, evidencing a significant difference across the play conditions.....	114
Table 11: Multiple paired-samples t-tests with a Bonferroni adjustment, evidencing statistically significant differences in means. ....	115
Table 12: Output of one-way within-subjects ANOVA for relative Beta frequency activity, evidencing a significant difference across Play conditions. ....	116

Table 13: Multiple paired-samples t-tests with a Bonferroni adjustment, evidencing statistically significant differences in means. ....	116
Table 14: Output of one-way within-subjects ANOVA for relative Gamma frequency activity, evidencing a significant difference across Play conditions. ....	118
Table 15: Multiple paired-samples t-tests with a Bonferroni adjustment, evidencing statistically significant differences in means. ....	118
Table 16: Summary of procedures used to register and enter events into the Catalogue of Events. ....	154
Table 17: Descriptive statistics for aggregate affect and arousal, averaged across the entire game play session for each participant. Data in % of maximum possible score. ....	163
Table 18: Results of Shapiro-Wilk test for normality, conducted for the aggregate arousal and valence scores. Both arousal and valence were found to be normally distributed.....	163
Table 19: Descriptive statistics for aggregate modelled emotional categories (fun, excitement, boredom, frustration, and challenge). Each of these emotional categories was averaged across the entire game play session for each participant. Data in % of maximum possible score. ....	165
Table 20: Results of Shapiro-Wilks test for normality, conducted for the aggregate fun, excitement, boredom, frustration and challenge scores. All modelled emotions were found to be normally distributed. ....	166
Table 21: Descriptive statistics for aggregate positive and negative facial expressions observed during game play. Data is mean expression duration in seconds. ....	167
Table 22: Descriptive statistics for modelled AV space and emotions, averaged across game event occurrences for each participant. Data in % of maximum possible score. ....	168
Table 23: Results of Shapiro-Wilk test for normality, conducted for modelled AV space and emotions. All data was found to be normally distributed.....	169
Table 24: Results of paired-sample t-tests comparing mean Speed Increase events against mean overall game experience. Mean differences and the significance of these are demonstrated for each of the five modelled emotions. ....	171



Table 25: Descriptive statistics for modelled AV space and emotions, averaged across each time non-useful Tetriminos were received for each participant. Data in % of maximum score.....	172
Table 26: Results of Shapiro-Wilk test for normality, conducted for modelled AV space and emotions. All data was found to be normally distributed.....	172
Table 27: Results of paired-sample t-tests comparing mean Non-Useful Tetriminos Received events against mean overall game experience. Mean differences and the significance of these are demonstrated for each of the five modelled emotions. ....	175
Table 28: Descriptive statistics for modelled AV space and emotions, averaged across each time Tetriminos were placed correctly for each participant. Data in % of maximum score.....	176
Table 29: Results of Shapiro-Wilk test for normality, conducted for modelled AV space and emotions. All data was found to be normally distributed.....	176
Table 30: Results of paired-sample t-tests comparing mean Correct Tetriminos Placement events against mean overall game experience. Mean differences and the significance of these are demonstrated for each of the five modelled emotions. ....	179
Table 31: Descriptive statistics for modelled AV space and emotions, averaged across each time Tetriminos were incorrectly placed for each participant. Data in % of maximum score.....	180
Table 32: Results of Shapiro-Wilk test for normality, conducted for modelled AV space and emotions. All data was found to be normally distributed.....	180
Table 34: Results of paired-sample t-tests comparing mean Incorrect Tetriminos Placement events against mean overall game experience. Mean differences and the significance of these are demonstrated for each of the five modelled emotions. ....	183
Table 35: Descriptive statistics for modelled AV space and emotions, averaged across each occurrence of a player clearing multiple lines for each participant. Data in % of maximum score.....	184
Table 36: Results of Shapiro-Wilk test for normality, conducted for modelled AV space and emotions. All data was found to be normally distributed.....	184

Table 37: Results of paired-sample t-tests comparing mean Multiple Line Clear (“Tetris”) events against mean overall game experience. Mean differences and the significance of these are demonstrated for each of the five modelled emotions. ...	187
Table 38: Descriptive statistics for modelled AV space and emotions, averaged across Player Win events for each participant. Data in % of maximum score. ....	188
Table 39: Results of Shapiro-Wilk test for normality, conducted for modelled AV space and emotions. All data was found to be normally distributed.....	188
Table 40: Results of paired-sample t-tests comparing mean Player Win events against mean overall game experience. Mean differences and the significance of these are demonstrated for each of the five modelled emotions. ....	191
Table 41: Descriptive statistics for modelled AV space and emotions, averaged across Player Loss events for each participant. Data in % of maximum score.....	192
Table 42: Results of Shapiro-Wilk test for normality, conducted for modelled AV space and emotions. All data was found to be normally distributed.....	192
Table 43: Results of paired-sample t-tests comparing mean Player Loss events against mean overall game experience. Mean differences and the significance of these are demonstrated for each of the five modelled emotions. ....	195
Table 44: Descriptive statistics for modelled AV space and emotions, averaged across occurrences of player opponents correctly placing Tetriminos for each participant. Data in % of maximum score. ....	196
Table 45: Results of Shapiro-Wilk test for normality, conducted for modelled AV space and emotions. All data was found to be normally distributed.....	197
Table 46: Results of paired-sample t-tests comparing mean Opponent Correct Tetriminos Placement events against mean overall game experience. Mean differences and the significance of these are demonstrated for each of the five modelled emotions. ....	199
Table 47: Descriptive statistics for modelled AV space and emotions, averaged per Opponent Incorrect Tetriminos Placement event. Data in % of maximum score. ....	200
Table 48 : Results of Shapiro-Wilk test for normality, conducted for modelled AV space and emotions. All data was found to be normally distributed.....	200

Table 49: Results of paired-sample t-tests comparing mean Opponent Incorrect Tetriminos Placement events against mean overall game experience. Mean differences and the significance of these are demonstrated for each of the five modelled emotions.....	203
Table 50: Descriptive statistics for modelled AV space and emotions, averaged across Opponent Multiple Line Clear (“Tetris”) events for each participant. Data in % of maximum score.....	204
Table 51: Results of Shapiro-Wilk test for normality, conducted for modelled AV space and emotions. All data was found to be normally distributed.....	204
Table 52: Results of paired-sample t-tests comparing mean Opponent Multiple Line Clear (“Tetris”) events against mean overall game experience. Mean differences and the significance of these are demonstrated for each of the five modelled emotions. ....	207
Table 53: Results of Shapiro-Wilk test for normality, conducted in this case for the Game Experience Questionnaire CORE components under the Equipment On experimental condition. All data was found to be normally distributed, including those for data not presented.....	212
Table 54: Descriptive statistics for Game Experience Questionnaire CORE components, averaged across all participants, and separated into Equipment On and Equipment Off conditions. Data in % of maximum score.....	213
Table 56: Results of paired-sample t-tests comparing mean subjective GEQ CORE components of game experience for Equipment On against Equipment Off conditions. Mean differences and the significance of these are demonstrated for each of the seven components.....	216
Table 57: Descriptive statistics for Game Experience Questionnaire SOCIAL components, averaged across all participants, and separated into Equipment On and Equipment Off conditions. Data in % of maximum score.....	217
Table 58: Results of paired-sample t-tests comparing mean subjective GEQ SOCIAL components of game experience for Equipment On against Equipment Off conditions. Mean differences and the significance of these are demonstrated for each of the three components. ....	218

Table 59: Results of Shapiro-Wilk test for normality, conducted for the Self-Assessment Manikin components for both Equipment On and Equipment Off experimental conditions. All data was found to be normally distributed.....	219
Table 60: Descriptive statistics for SAM components, averaged across all participants, and separated into Equipment On and Equipment Off conditions. Data in % of maximum score.....	222
Table 62: Results of paired-sample t-tests comparing mean subjective SAM components of affective experience for Equipment On against Equipment Off conditions. Mean differences and the significance of these are demonstrated for each of the three components. ....	223
Table 63: Descriptive statistics for observed facial expressions, averaged across all participants, and separated into Positive and Negative for both Equipment On and Equipment Off conditions. Data in seconds of expression duration. ....	224
Table 64: Results of paired-sample t-tests comparing mean observed Positive and Negative Facial Expressions for Equipment On against Equipment Off conditions. Mean differences and the significance of these are demonstrated for each of the three components. ....	225
Table 65: Descriptive statistics comparing Subjective and Objective measures of Valence and Arousal, averaged across all participants, and separated into Equipment On and Equipment Off conditions. Data in % of maximum score. ....	226
Table 66: Pearson's test for correlation between Subjective and Objective measures of Arousal, demonstrating a strong positive relationship. ....	229
Table 67: Pearson's test for correlation between Subjective and Objective measures of Valence, demonstrating a medium positive relationship. ....	229

## Chapter 1: Introduction

### 1.1. Project Background

The aim of this study is to investigate the events occurring during video game play that can nuance the overall emotional experiences of players. Specifically, while previous studies largely examine overall player experiences, and while some have attempted to examine responses to specific game events, an examination into how these game events occurring during game play contribute towards and compare against the overall gameplay experience is not sufficiently understood.

This research project undertakes a novel exploratory experiment to identify ten game events occurring during video game play that elicit unique affective responses, and identifies how these contribute towards the overall emotional experience of game play. These ten game events analysed against five modelled emotions, comparing each against the average game experience. Additionally, the influences of measurement effects upon game play experiences are examined within the context of video game research, identifying effects such as significant changes in self-reported levels of immersion, and awareness of both positive and negative emotional experiences. These changes could influence the interpretation of psychophysiological data, and provide valuable insights into the use of physiological measures for the purposes of player experience evaluation.

The first chapter of this thesis provides an overview of the research context, and places this project within its field of study. Specifically, it provides an overview of the purpose of undertaking this study, provides a description of the research questions posed, and the research strategy employed. This chapter then presents an overview of the context into which this study can be placed, a summary of the contributions of the research, and a chapter guide detailing the structure for the remainder of the thesis.

### 1.2. Purpose of this Research Project

As will be discussed in detail in the following chapter, due to gaps in current understandings of real-time responses to game events and methods capable of measuring them, an exploratory investigation into the affective responses elicited through events occurring during video game play was undertaken using psychophysiological objective measures of player emotions. More specifically, the research was designed to test a method capable of investigating the relationship between game events occurring in real-time during game play, and the affective responses that players experience when undertaking them. This includes an examination of how the process of measuring player emotions through psychophysiological instruments influences the overall game play experience.

This research was initially conceptualised entirely within the domain of video game research, intending to explore the influence that individual components of video games elicited upon the overall game play experience, and explore how these might be better understood for the purposes of constructing serious games. However, the issue of how to reliably measure and assess game components and player experiences as they exist during game play led to the secondary domain of psychophysiology.

The results of research are only valuable if they elucidate previously unknown phenomena, or contribute improved insights and understandings for those phenomena that have not previously been well explained. This research therefore operates within the bounds of two different research fields – video game research, and psychophysiology research. While both areas represent diverse fields with rich histories, the junction of these two domains provides an opportunity within which this research project can be placed to explore video game experiences in real-time.

As will be discussed in detail in the next chapter, this research is specifically focused on two issues:

- Measurement of affective responses to specific events occurring in real-time during game play; and
- Measurement effects upon game play experiences elicited through the use of psychophysiological equipment.

While video games and the experiences that they elicit have received a lot of academic attention (see Chapter 2 for discussion), responses to events occurring during or within video game play are not well understood due to the complex nature of video game experiences (with a wide range of overlapping mechanics and confounding factors) and the historically limited temporal-resolution research methods used to understand emotions and experiences.

Historically, studies using subjective methods often attempted to understand player experiences through reductive techniques, where the researcher attempts to control as many factors as possible and manipulate only those of interest in an attempt to elucidate insights into the influence of these factors. While many significant insights can and have been developed through use of such methods, there are a number of problems associated with these methods (see Chapter 2 for discussion), including:

- Limited ability to collect data on player experiences as they occur during game play (methods such as ‘think-aloud’ can be useful, but risk changing the play experience due to breaking immersion, amongst other issues); and
- Reducing the game components to test limited features or events alters the video game experience from what it might usually be.

This project wished to investigate player emotions and affective responses as they occur during game play, and ideally within a natural game play environment that has not been simplified or manipulated. It was therefore the intent of this project to attempt a different approach through which video game experiences and affective responses therein can be examined.

The juncture of the video game research domain with the domain of psychophysiology is not a unique one, with a number of researchers using psychophysiology to examine video games over the past decade. Psychophysiology, which is discussed in detail in Chapter 2, uses sensors to collect objective data from physiological signals such as heart rate, and uses these to infer psychological processes and states (Cacioppo, Tassinary and Berntson, 2007; Cowley *et al.*, 2016). Psychophysiology has a number of benefits that can provide insights into real-time events, including those of interest to this project:

- High-temporal resolution, potentially providing thousands of discrete measures per second in real time, during video game play;
- Covert measures that do not require participants to interpret questions correctly; and
- Facilitates examination of game events without requiring the alteration or simplification of video games.

While psychophysiology provides an opportunity, with several benefits that can be leveraged for the purposes of this research, there are several limitations, including:

- Psychological inferences are not able to be perfectly mapped or modelled using physiological measures, due to complex physiological processes having competing and potentially confounding influences (Nacke, 2013; Nacke, 2015).
- Complexity of data, often requiring substantial data processing and analysis in order to produce meaningful insights.

This provides an opportunity whereby the affective responses to events occurring within video games are able to be examined at high-temporal resolutions without requiring the alteration or manipulation of video games to do so. Specifically, this research aims to identify video game events that occur during game play, and analyse differences in affective responses that are elicited through interaction with these.



### 1.3. Research Questions and Strategy

To provide insights into the above-mentioned areas of interest within the domains of video game research and psychophysiology, it was necessary to scope the investigation such that it addressed a specific problem, with research questions appropriate for addressing this problem. As will be discussed in the next chapter, the area of concern of this thesis focused on the apparent lack of attention paid to affective responses as they occur during video game play. One potential solution to addressing this has been to use psychophysiology, which offers continuous and non-disruptive measures, to facilitate in-depth explorations of affective responses and the mechanisms through which they can be elicited.

#### 1.3.1. Research Questions

With respect to the above-mentioned area, a general research problem was formulated as:

**RP:** *How can affective responses to real-time video game events be objectively measured using a novel context-rich psychophysiological method?*

To address this problem, two dimensions of the research problem were discovered, resulting in the following research questions:

**RQ1:** *How do events occurring during video game play influence event-related affective responses?*

And

**RQ2:** *How does use of psychophysiological instruments influence the player experience during video game play?*

As a result of the above two research questions, the goals of this research were to undertake an exploratory study to collect data on game events occurring during video game play; and undertake an analysis of the affective responses elicited through interaction with these events.

### 1.3.2. Research Methods for this Thesis

This project employs a new adaptation of a psychophysiological approach to investigating affective responses, which has enabled the examination of game experiences at multiple levels of granularity in real-time. Selection of research methods used in this research project was undertaken on the basis of their appropriateness in fulfilling the above-mentioned objectives of the research. As fully explicated in Chapter 3, this research used multiple methods and data collection instruments, which were developed and tested through a series of pilot studies.

The foremost method, however, was the undertaking of an experiment using objective psychophysiological equipment and techniques, also incorporating multiple subjective techniques. Both objective and subjective analysis techniques were undertaken to analyse the research data collected within the experimental environment.

### 1.4. Context of the Research

To provide insight into the research questions previously stated, this project undertakes experimental data collection on members of the Tasmanian public. As will be discussed in Chapter 3, one video game was used for experimental purposes, namely Tetris Party Deluxe on the Nintendo Wii platform, as this is a relatively simple but popular video game from which insights into affective responses could be drawn.

This study includes the examination of ten events observed to occur repeatedly during play sessions of Tetris Party Deluxe, and analyses affective responses to these based upon five modelled emotions as classified using a fuzzy logic system.

### 1.5. Summary of Contributions

This research project makes several distinct contributions to theoretical knowledge and the research community for the domains of both video game and psychophysiology research:

- This research will present a novel adaptation of existing psychophysiological methods to measure and analyse event-related affective responses to observed video game events. Ten events were identified during video game play and analysed against five modelled emotions, with differences from the mean affective experience examined for each. These revealed insights into the experiences and game mechanics underlying these affective responses, as well as the events that elicited them.
- This research further examined incidences of events where the event occurred for the player, and compared these against incidences where the same type of event occurred for their opponent. This provided novel insights into the interplay between the personal experience of events and events occurring to other players.
- This research examined measurement effects associated with the use of psychophysiological equipment, revealing that the presence of equipment significantly reduced self-reported immersion, and significantly increased self-reported challenge, as well as perceptions of both positive and negative emotions. This provided novel insights into measurement effects elicited through use of psychophysiological equipment, with indications that participants became distracted and introspective when electrodes were placed on their bodies.

### 1.6. Layout of Chapters

The following sections briefly describe the layout of this thesis, and the main contribution of each chapter.

#### 1.6.1. Chapter 1: Introduction

*Chapter 1: Introduction* has provided insight into the research project, has briefly identified the research area, and the associated questions to be addressed within the project. Specifically, this chapter has given an outline of the purpose of this project, and provided an indication of the contexts within which this study and its results may be placed.

#### 1.6.2. Chapter 2: Literature Review

*Chapter 2: Literature Review* provides an overview of existing research and literature within the domains of both video game research and psychophysiology research, with a specific focus upon emotions and experiences of video game play. This chapter identifies that although the ability of video games to elicit emotional responses and experiences, understandings of the events and mechanics contributing towards these experiences are underexamined.

Specifically, this chapter identifies the theoretical environment leading to the commencement of this research project.

#### 1.6.3. Chapter 3: Methodology

*Chapter 3: Methodology* details the research process that was undertaken throughout this project. It discusses, both theoretically and practically, the philosophical assumptions underpinning the research, the processes, and procedures used throughout the project. This includes a discussion of the pilot studies undertaken for the purposes of improving the study design, as well as a discussion of the techniques used to identify game events, and model emotions from psychophysiological data.

Specifically, this chapter identifies the nature of the research undertaken, the overarching methodology adopted, and the specific research and analytical techniques used within this study.

### 1.6.4. Chapter 4: Results

*Chapter 4: Results* provides details of statistical tests undertaken to explore the data collected through experimentation. Specifically, statistical results are provided for each of the ten observed game events, alongside results for overall game play experiences, correlations between subjective and objective measures, and for measurement effects upon game play experiences. This chapter therefore provides insights into the research questions stated in Section 1.3.1.

### 1.6.5. Chapter 5: Discussions

*Chapter 5: Discussions* provides an examination of the results derived through this project, while introducing academic literature in order to draw insights with regard to the interpretation and implications of the research findings.

### 1.6.6. Chapter 6: Research Summary

*Chapter 6: Research Summary* provides the concluding remarks of the thesis, and provides an overview of the implications that can be drawn from the research project. An analysis of the research project as a whole is also provided.

## 1.7. Chapter Reflections

This research project involves an exploration of video games and the affective experiences that may be derived from them, with a specific focus upon how events occurring during video game play can elicit specific affective responses that nuance the overall game play experience.

Despite the rapid growth of video games as a form of entertainment, the emotional correlates of game play experiences are poorly understood. This study was therefore designed to identify game play events that contribute towards emotional experiences, and assess the affective responses elicited through interaction with these events.

Having established a preliminary insight into the objectives and background of this research project, the following chapter will provide a detailed discussion of the context within which this study can be placed, and the theoretical environment within which this study has been conducted.

## Chapter 2: Literature Review

### 2.1. Chapter Introduction

This chapter provides an overview of the theoretical framework in which this thesis can be placed. Specifically, this chapter has been divided into two main sections, each providing theoretical insight into the development of the research context. The first section introduces video games and the history of research within this domain, with a specific focus upon examination of emotions and player experiences. The second section then introduces the domain of psychophysiology, including an introduction to physiological signals, and how these can be used to infer affect, as well as an exploration of existing studies using psychophysiology within video game contexts.

### 2.2. Domain 1: Video Game Research

The value in developing our understandings of video games and the influence that video game components have on individuals is more important than ever. Video

games have reached an unprecedented level of ubiquity within Australia, with over 93% of Australian households contain at least one device that is used for playing video games on a daily basis, over 76% of video gamers being over the age of 18, 19% of video gamers are over the age of 51, and the average of video gamers being 32 (Brand, Lorentz and Mathew, 2014).

Australians in their 40's and 50's now comprise the fastest growing group of video gamers, the majority of whom report that they play video games to keep their minds active, which they perceive to be of increasing importance for their ongoing health and wellbeing (Brand, Lorentz and Mathew, 2014). Furthermore, while video games have traditionally been perceived as a male dominated activity, the proportion of males to females who regularly play video games is now nearly identical, with 47% of video gamers being female (Brand, Lorentz and Mathew, 2014). Given that the average Australian is 37 years old and that the Australian population has roughly equal numbers of males and females (ABS, 2013), it is evident that video games have become a normalised and significant medium of entertainment that is widely enjoyed by the majority of the Australian population.

Video games now represent a significant economic and social influence within Australia, with the Australian population spending an estimated 7.8 billion hours playing video games per year, and spending an average of over \$2.3 billion per year on video games and accessories (Brand, Lorentz and Mathew, 2014). The video game market in Australia has seen an economic growth of approximately 8% (Compound Annual Growth Rate) per year since 2010, which is expected to continue for at least the next 3 years (Brand, Lorentz and Mathew, 2014). To put this in perspective, this means that video games now represent at least 10% of the \$22.5 billion (Ma, 2014) that Australians spend on entertainment per year. When considering that this means that the average Australian is spending approximately 0.5% of their total annual income on video games (ABS, 2013; Ma, 2014), and spending approximately 337 hours per year playing video games (Brand, Lorentz and Mathew, 2014), the social and economic impacts of video games within Australia are evidently significant.

Given the ubiquity that video games enjoy both within Australia as well as other countries, it is not surprising that the study of video games is a rapidly growing field of research globally. Research in the video game field now include a broad range of areas, ranging from the improved design of entertainment video games, to using video games for serious purposes in areas such as education and training, health care, energy conservation, change management, and marketing.

### 2.2.1. Social Video Game Play Experiences

While all video games are generally highly engaging and fun activities, there has been a recent shift towards social video games, with the majority of games being released now including variety of social features (Kirman, 2010; Kirman *et al.*, 2011). Recent studies have found that playing video games socially results in significantly higher levels of player engagement and more positive and arousing emotional experiences when compared to non-social games (Lim and Reeves, 2010; Gajadhara *et al.*, 2008; Clark, 2012; Klarkowski *et al.*, 2016). Player purchasing behaviours appear to also support this trend, with 71% of Australian video gamers report that social interactions are among the most important aspects of video games (Brand, Lorentz and Mathew, 2014).

#### 2.2.1.1. Types of Social Interaction

Among the most significant factors to consider in social video games is the type of social interactions facilitated, which can be divided into those that facilitate competitive interactions and those that facilitate cooperative interactions. Within these types of social interactions, games can include different mechanics that control the specific ways in which players are able to interact, which can be divided into those that facilitate direct interactions and those that facilitate indirect interactions. Social interactions within video games have traditionally been competitive in nature, however the last decade has seen the extensive and sweeping



addition of cooperative interactions to many video games (El-Nasr *et al.*, 2010; Lim and Reeves, 2008; Kirman, 2010; Kirman *et al.*, 2011).

**Competitive interactions:** Interactions where two or more players are engaged as adversaries within the video game environment, striving to win by defeating the opposition (Decety *et al.*, 2004).

- **Direct mechanics:** Adversaries can directly influence each other within the game, with a direct impact upon their respective performances. For example, in a racing game one player might bump another player off the track, and thus gain an advantage by negatively influencing the performance of the other player.
- **Indirect mechanics:** Adversaries cannot directly influence each other within the game, but can monitor their own performance and compare it to the performances of other players. For example, in a puzzle game one player might be able to see the scores of other players, which will indirectly influence their performance as the player strives to beat the score of the other players.

**Cooperative interactions:** Interactions where two or more players are engaged as allies within the video game environment, striving to achieve a mutual goal (Lim and Reeves, 2008).

- **Direct mechanics:** Allies can directly influence each other within the game, with a direct impact upon the combined performance of their team. For example, in many role-playing games allies will need to work together and support one another in order to succeed, and thus mutually influence the performance of their team.
- **Indirect mechanics:** Allies cannot directly influence each other within the game, but can monitor their respective performances and communicate to assist each other indirectly. For example, in some puzzle and shooter games allies can watch one another as they play and provide advice on how they could handle situations or improve their performance.

### 2.2.1.2. *Types of Social Contexts*

An additional factor to consider in the investigation of the affective nature of social games is the social environment within which the video games are played, which have been found capable of having a significant influence upon the experience when playing video games (Ballard, Visser and Jocoy, 2012). Assuming that the social context is as naturalistic as possible (i.e. socializing with friends within a comfortable environment), these can be divided into local and remote contexts.

- **Co-located Players:** Social contexts or environments where the players are situated within the one space, and can directly communicate without the assistance of technology. For example, a group of friends are playing a video game together within the one room and are interacting socially, both internally through game mechanics and externally through bodily communication. Recent studies have demonstrated that playing socially with co-located players elicits significantly stronger arousal and game enjoyment when compared against remote players, or non-player characters (Mandryk and Inkpen, 2004; Nacke, Stellmach and Lindley, 2011; Clark, 2012).
- **Remote Players:** Social contexts or environments where the players are situated in geographically remote spaces, and are not able to directly communicate without the assistance of technology. For example, two friends who live in different buildings are playing a video game together online and are interacting socially within the game through game mechanics, and may be communicating verbally with the assistance of technology.

### 2.2.2. Existing Measurement of Game Experiences

The study of emotions and behaviours within video games has been traditionally dominated by subjective and qualitative measures, typically relying upon self-report surveys, interviews, and observations (Nacke, 2009; Mandryk, Atkins, and Inkpen, 2006). While these can be useful instruments that can provide insights into the

perceptions and extrinsic behaviours of people, they are subject to response bias, varying interpretations of survey questions, poor recollection, observational bias, and are generally unable to provide data on the factors that are underlying their measures (Klasen *et al.*, 2008; Mandryk and Atkins, 2007; Cowley *et al.*, 2016).

Additionally, subjective measures are typically measured after a participant has completed an activity or require interrupting the participant during the activity, and so cannot measure time-sensitive responses to stimuli in a naturalistic and non-invasive manner. One form of objective measurement that has been attempted is video recording participants as they play and undertaking observational analysis, however this can suffer from researcher misinterpretation, and requires an extremely time-consuming codification of the video data (Mandryk and Atkins, 2007).

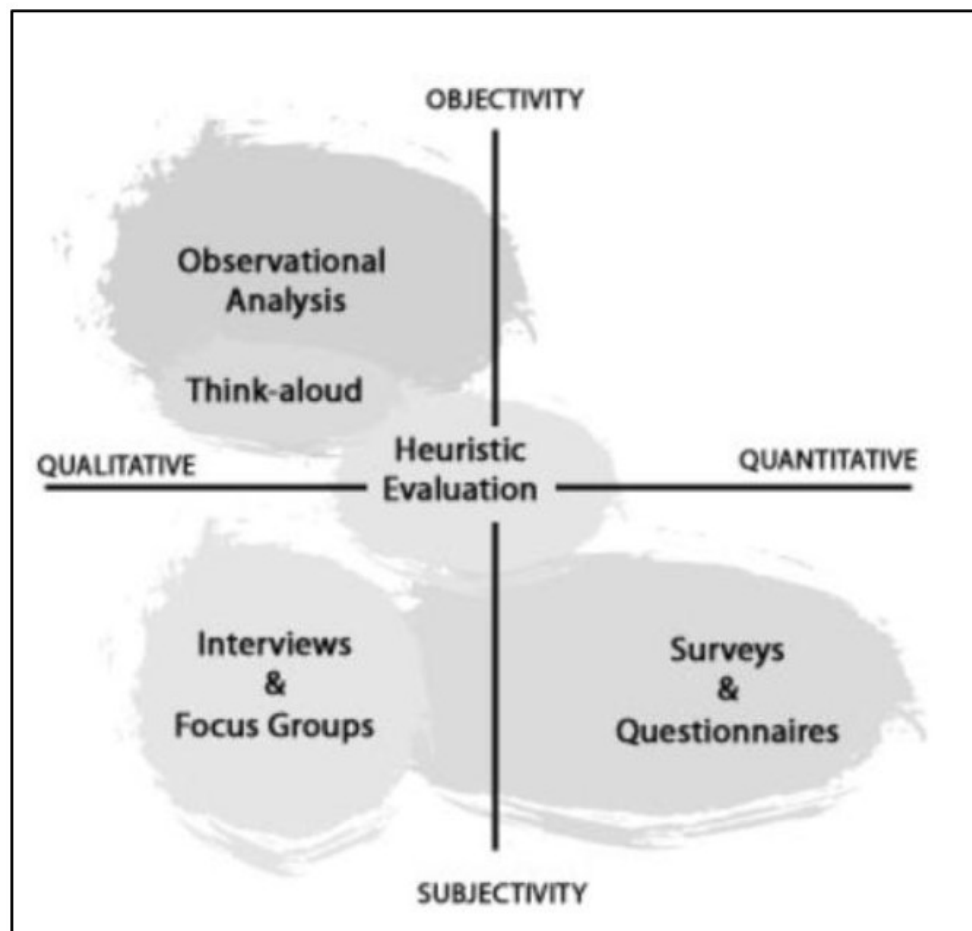


Figure 1: The use of traditionally subjective and qualitative instruments within studies of video games, with a notable gap in instruments that are both objective and quantitative (Nacke, 2009).

A recent and rapidly growing interdisciplinary science that attempts to avoid these issues is psychophysiology, which involves objectively measuring quantitative aspects of the human body in order to assess bodily changes. These measures are then used to understand the underlying processes of the human brain and the central nervous system, particularly as they relate to the experience of emotions and behaviour (Mandryk & Nacke, 2016; Nacke, 2015; Nacke, 2013; Cacioppo, Tassinary and Berntson, 2007).

### 2.2.3. Shift towards Psychophysiology

More recently there have been an increasing number of studies using psychophysiology, which involves the examination of bodily changes and signals, to understand neural events and the associated psychological processes that are underlying these changes (Nacke, 2009; Cacioppo, Tassinary and Berntson, 2007).

Studies utilising psychophysiology have examined a range of phenomena within video games (for an overview, see Mandryk & Nacke, 2016; Kivikangas et al., 2011), including emotional and cognitive experiences (e.g. Nacke, Stellmach and Lindley, 2011), video game addiction and problematic behaviours associated with addiction (e.g. Sim *et al.*, 2012), and correlations between video game play and violence (e.g. Weber, Ritterfeld and Mathiak, 2006; Ashbarry *et al.*, 2016). These studies have resulting in several profound and significant findings, including links between video games and pro-social behaviours (Greitemeyer and Osswald, 2010), anti-social behaviours (Weber, Ritterfeld and Mathiak, 2006), improved cognitive and visual attention (Gago and Almeida, 2013), increased positive emotions (Salminen and Ravaja, 2007), and decreased stress (Russell and Newton, 2008), to name a few.

The following section provides a discussion of psychophysiology as it exists within its own respective domain, followed by a discussion of how the domains of video games research psychophysiology research have been used together to improve the measurement of phenomena of interest to video games researchers.

### 2.3. Domain 2: Psychophysiology

#### 2.3.1. Overview

Broadly speaking, psychophysiology is the use of physiological signals to study and make inferences about psychological phenomena, under the assumption that the human body is the medium through which all things are experienced (Cacioppo, Tassinary, and Berntson, 2007; Kivikangas *et al.*, 2011). This allows psychophysiology to measure objective data on processes and phenomena that are inherently subjective in nature, providing valuable insights into psychological processes (Kivikangas *et al.*, 2011).

Psychophysiology avoids many of the pitfalls of subjective measures due to its very nature, as it collects objective-quantitative data through passive and covert electrodes, which collect electrical signals emitted by the human body (IJsselsteijn *et al.*, 2008). This form of measurement can thereby avoid response bias, attempts at deception, and does not require participants to interpret survey questions in a consistent manner (Drachen *et al.*, 2010). These benefits have resulted in a recent growth in adoption of psychophysiological instruments within studies of video game experiences, as well as many other fields of research (Mandryk & Nacke, 2016; Nacke, 2015; Ambinder, 2011).

#### 2.3.2. Psychophysiological Signals

Psychophysiology involves the collection of objective data through physiological signals, which are measured via electrodes connected to biometric instruments, in order to provide data on the affective state of an individual, as well as neural activities and processes (Cacioppo, Tassinary and Berntson, 2007; Nacke, Drachen

and Gobel, 2010). There are a wide range of psychophysiological signals that can be used, measuring dozens of different aspects of the human body. For the sake of brevity, only those of direct interest to this project are discussed.

### **2.3.2.1. Brain Activity:**

Brain activity is, at its simplest definition, the movement of electricity between brain cells as they communicate with one another (Cacioppo, Tassinary and Berntson, 2007). These electrical movements are responsible for every thought, every memory, and every action that humans make (Sukel, 2011; Cacioppo, Tassinary and Berntson, 2007). Despite the best efforts of researchers for centuries, much of the brain and how it works remains unknown (Sukel, 2011; Cacioppo, Tassinary and Berntson, 2007). The development of new technologies has improved the ability of researchers to measure different aspects of brain activity, and it is now possible to measure and observe electrical activity as it occurs throughout the brain (Cacioppo, Tassinary and Berntson, 2007).

It is now possible to measure brain activity while people are engaging with activities and experiencing stimuli, which over time this has allowed for varying types and regions of brain activities to be associated with specific functions, processes, and mental states. A summary of these can be seen in Figure 3, which highlights the core brain structures and their associated general functions.

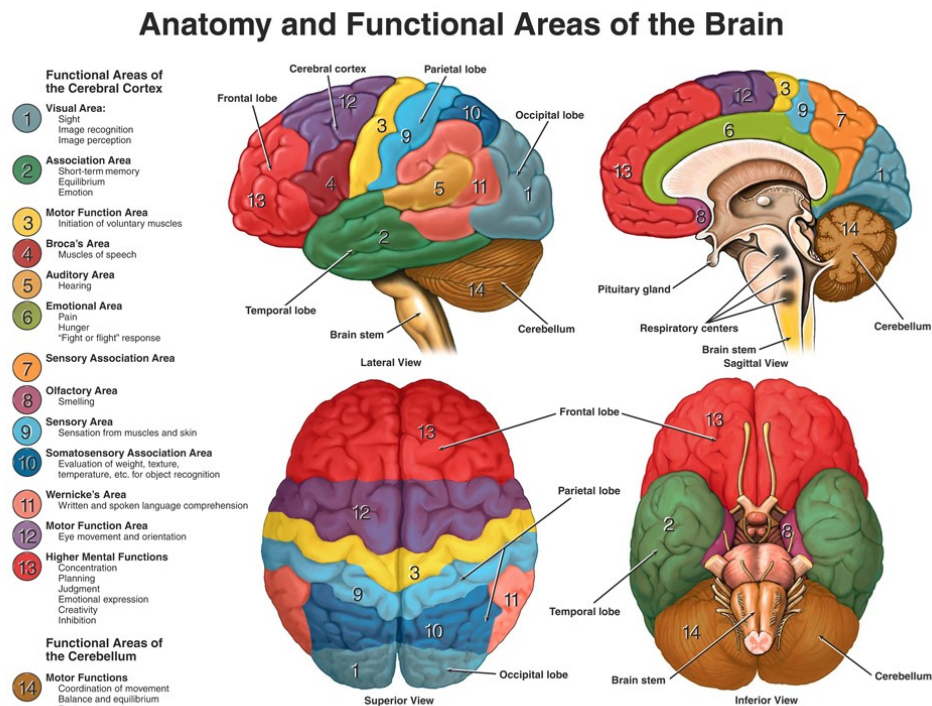


Figure 2: Functional brain regions and their associated mental activities and processes (Sukel, 2011).

The ability to measure brain activities is increasingly important within the field of video games, as video games involve people interacting with engaging and often emotional virtual experiences, which necessarily involves a large number of mental activities and processes (Dolcos, Iordan and Dolcos, 2011; Han *et al.*, 2010). A number of studies have successfully measured brain activity during video game play, and have developed significant insights into mental processes, including increased activity in regions associated with visual spatial processing (Havranek *et al.*, 2012), increased activity in regions associated with learning and information processing (Dye, Green and Bavelier, 2009), and increased activity in regions associated with motor functions (Klasen *et al.*, 2012). Whilst the discussed extant research has demonstrated that video games can have profound affective influences upon players, the specific mechanisms in which these influences occur is not known.

### Electroencephalography (EEG)

Electroencephalography (EEG) records the electrical currents discharged when neurons within the brain are active (Nacke, Stellmach and Lindley, 2011). The level of neuronal activity within a specific region of the brain is indicative of how activate that brain region is. For example, if the electrical discharge within a region is increasing then this would indicate a higher level of brain activity within that region, and if the electrical discharge within a region is decreased then this would indicate a lower level of brain activity within that region (Nacke, Stellmach and Lindley, 2011). These electrical discharges also have different electrical frequencies, with different bands of frequencies being associated with different brain activities or states (Wehbe & Nacke, 2013; Lim and Reeves, 2010).

Measuring the electrical discharges from neuronal activity becomes valuable when considering the specific frequency bands that are being produced, as well as identifying the specific brain regions that these are being produced in. Different frequency bands are associated with varying mental states, and likewise different brain regions are associated with varying mental activities or processes. Previous studies of general play experiences using EEG have correlated psychophysiological arousal and engagement with specific brain activity patterns, which predominately occur within the beta and gamma frequency bands (Nacke, 2010; Nacke *et al.*, 2010; Nacke *et al.*, 2009; Ravaja *et al.*, 2005). The alpha frequency band has also been associated with brain activities that may be of interest to this project, namely that alpha frequency activity may be indicative of relaxation and mental workloads.

The brain activity frequency bands and their associated mental states (Nacke *et al.*, 2009; Lim and Reeves, 2010):

- Delta Band (0.1 - 3 Hz): Associated with deep sleep or meditative trances.
- Theta Band (4 - 7 Hz): Associated with being drowsy, inactive, or idle mental workloads.
- Alpha Band (8 - 12 Hz): Associated with being calm or relaxed, or moderate mental workloads.
- Beta Band (12 - 30 Hz): Associated with being alert, engaged, and focused.



- Gamma Band (35 - 45 Hz): Associated with actively processing information and intense mental workloads.

Figure 3 illustrates what raw EEG data can look when measured like across multiple frequencies. The blue signal at the top is the EEG signal in general, whilst the other coloured signals below it are the breakdowns for each of the frequency bands that comprise the overall EEG signal. EEG data is typically analysed statistically, through the aggregation of signal strengths across multiple participants, and then assessing the relative difference in frequencies and activation patterns across the dataset.

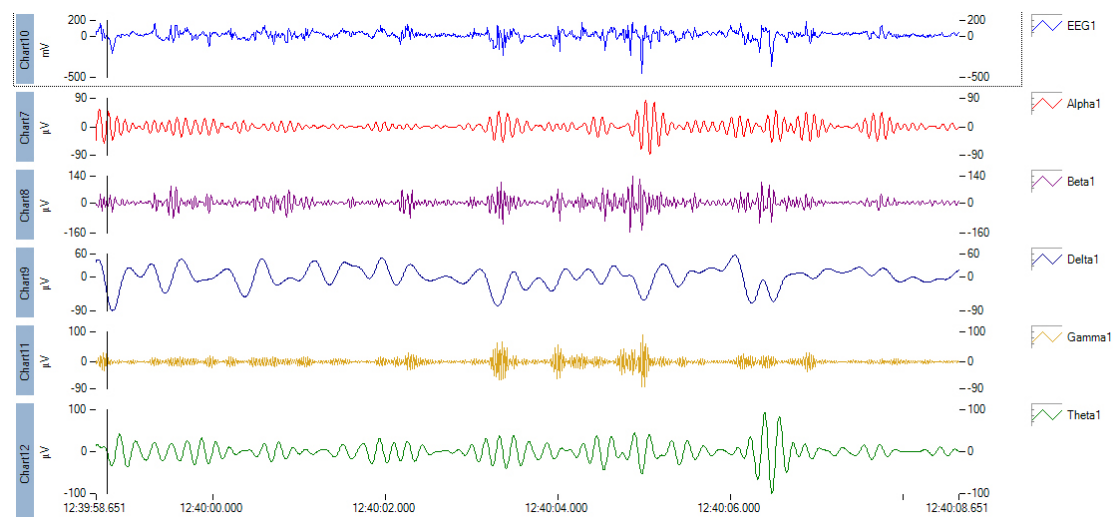


Figure 3: Raw EEG signals for multiple channels and frequencies, as recorded using the BioRadio device and software (GLNeuroTechnologies, 2014).

Figure 4 shows an alternative tool to assist in the analysis and interpretation of EEG data, where EEG signals are processed onto a heat map that illustrates the activation (red colours) and deactivation (blue colours) of different regions of the brain. Both forms of assessing EEG data allow researchers to assess brain activity frequencies and regions of activation in response to different stimuli and events, providing valuable insights into both cognitive and affective processes within the brain of research participants (Cacioppo, Tassinary and Berntson, 2007).

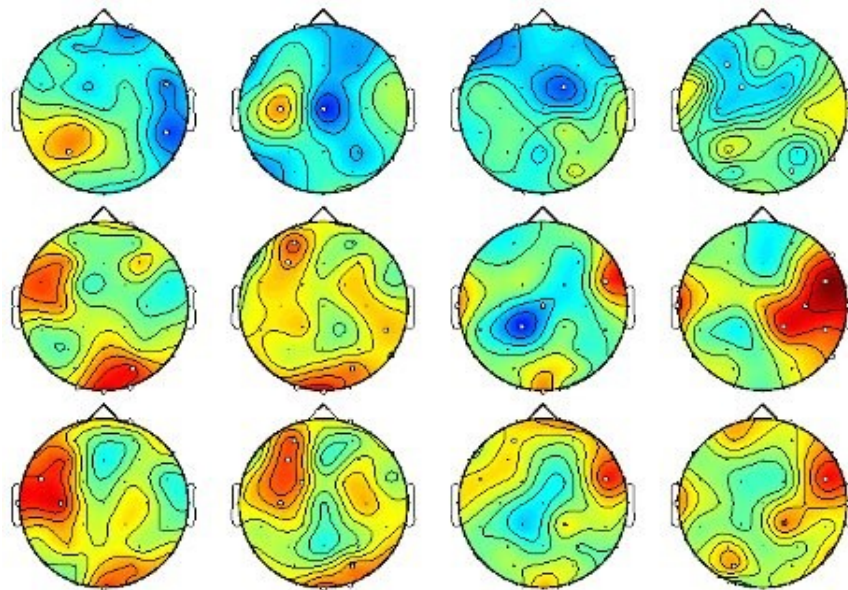


Figure 4: EEG heat-maps showing brain region activation and deactivation, created through processing the raw EEG signal data (Koelstra and Muhl, 2011).

#### 2.3.2.1.1. EEG Summary

EEG has a number of advantages when compared against alternative methods of measuring brain activity such as fMRI, which make it an ideal measure for affective responses to video games:

##### Benefits:

- **High temporal resolution:** EEG devices actively capture data during game play, and have a temporal resolution that allows for responses to be analysed at the millisecond level (Seo, Lee and Yoo, 2013).
- **Cost effective:** EEG devices, while still expensive, cost substantially less to purchase and operate when compared to alternatives such as fMRI. Lower quality devices may be purchased inexpensively, though the quality of the data output may suffer as a result.
- **Relative comfort:** EEG devices are could be setup and used within relatively comfortable and naturalistic environments. While movement and social

interactions will still result in data loss, it is possible that valuable data could still be captured during social video game play.

However, there are a number of limitations to EEG that must be considered in the design of the research:

### Limitations:

- **No measure for valence:** Although EEG is able to measure the arousal of participants, it is not easily able to provide a measure of the valence of this arousal due to currently limited understandings of the human brain (Nacke, Stellmach and Lindley, 2011). Valence must therefore be measured through alternative measures.
- **Artefacts:** Muscle movements can produce artefacts within EEG data, which necessitates the cleaning and processing of the data prior to analysis (Nacke, Stellmach and Lindley, 2011).
- **Low spatial resolution:** The spatial resolution of EEG is limited by the number of electrodes used, and some of the smaller brain regions can become hard to distinguish if a low number of electrodes are used (Havranek *et al.*, 2012).
  - EEG also only shows brain activity in the outer extremities of the brain, in those regions that are located in proximity to the scalp, and therefore is unable to measure brain activity in interior regions of the brain (Nacke, Stellmach and Lindley, 2011).

### 2.3.2.2. Heart Rate:

The heart is responsible for pumping blood throughout the body via a series of blood vessels to provide other organs with oxygen and nutrients, and to remove waste products (Hall, 2011; Cacioppo, Tassinary and Berntson, 2007). Depending on the requirements of the body, the rate at which the heart beats can vary. This can include increasing the supply of oxygen to other organs, or preparing the body for an expected sudden increase in oxygen requirements (Klabunde, 2013; Cacioppo, Tassinary and Berntson, 2007).

Heart rate variations that result from an expectation of increased oxygen requirements typically occur due to activation of the sympathetic nervous system, which is responsible for regulating the fight-or-flight response (Cacioppo, Tassinari and Berntson, 2007). The fight-or-flight response involves a series of physiological changes that occur in response to stimuli that are perceived as threatening or potentially harmful, which includes releasing adrenaline into the bloodstream, activating sweat glands, and increasing the heart rate in anticipation (Cacioppo, Tassinari and Berntson, 2007).

The sympathetic nervous system also influences the heart rate in response to other stimuli that are not necessarily threatening, but are rather exciting or enjoyable (Cacioppo, Tassinari and Berntson, 2007). In the case of video games, studies have shown that video games are capable of eliciting heart rate variations through the presentation of threatening and exciting stimuli (Drachen, Nacke and Yannakakis, 2010).

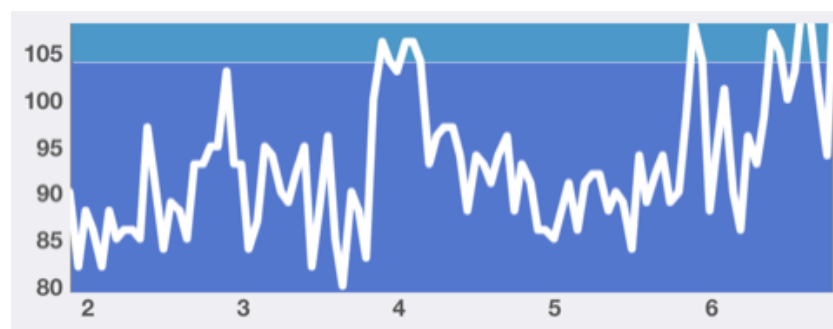


Figure 5: *Heart rate variations in response to video game stimuli, showing peaks and dips in the heart rate as different components within the game are experienced. Graph is an excerpt of data collected from one participant during pre-pilot stage of this research project.*

Heart rate variations also occur as a result of an increased demand for oxygen to organs, which occurs through the activation of the parasympathetic nervous system, which is responsible for regulating unconscious actions (Cacioppo, Tassinari and

Berntson, 2007). This will typically occur where the body is undertaking increased physical activity, and so the heart rate will increase in order to provide the muscles with the increased oxygen that they require to move (Cacioppo, Tassinary and Berntson, 2007). These parasympathetic heart rate variations are not of interest for this project, as video games do not typically require increases in physical activity. However, these influences must be considered within research design, as they could confound the measure of sympathetic heart rate variations that are of interest.

Increasing heart rate due to the activation of the sympathetic nervous system is a major indicator of physiological arousal, and can be used as a measure to determine the extent of affective arousal (Cacioppo, Tassinary and Berntson, 2007). The measurement of heart rate variations is the most commonly used physiological signal within extant psychophysiological video games research, and has been successfully used as a measure of arousal within a number of studies (e.g. Cowley *et al.*, 2014; Koelstra and Muhl, 2012; Drachen, Nacke and Yannakakis, 2010; Ravaja *et al.*, 2006). This makes the measurement of the heart rate an important consideration for this project, as it provides a simple objective measure of the affective influence that video game components can have.

### 2.3.2.2.1. Electrocardiogram (ECG)

The electrocardiogram (ECG) measures the electrical activity generated by the heart muscle as it contracts during heart beats, which can provide detailed data on the specific contraction patterns of the heart (HR), and the intervals between beats (IBI). The ECG collects electrical impulses that are generated through activation and deactivation of the muscles within the heart, which it uses to determine the wavelength and rate at which the heart is beating (Cacioppo, Tassinary and Berntson, 2007).

Higher quality ECG devices within clinical settings are also capable of identifying the size and position of the chambers within the heart, can detect damage to the heart, and can assess the influence of other factors such as drugs or pacemakers (Cacioppo,

Tassinary and Berntson, 2007). Depending on the type of ECG device used, between 3 and 12 electrodes, or leads, are used in varying locations on the body (Cacioppo, Tassinary and Berntson, 2007). Figure 6 below shows the correct electrode placement for a 7-lead device.

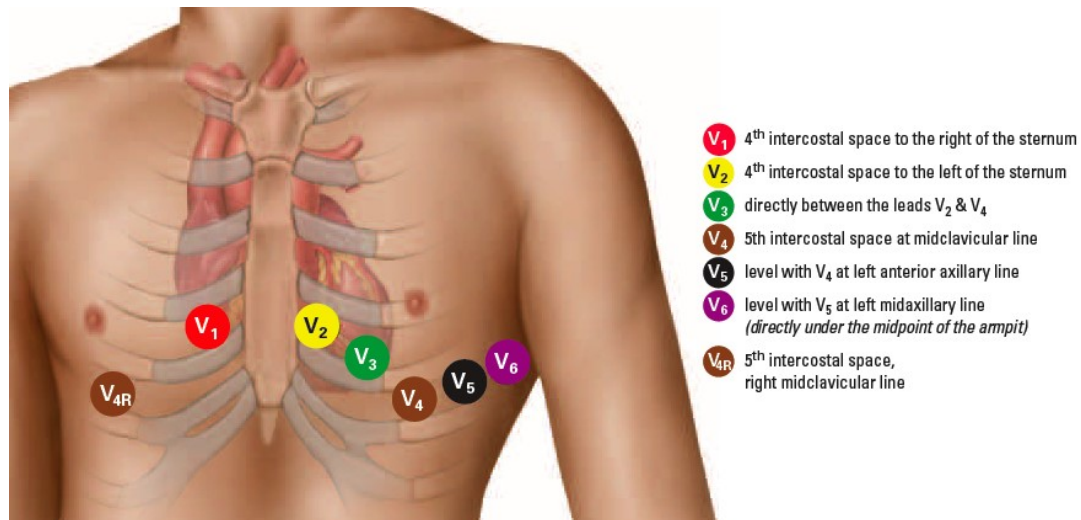


Figure 6: Diagram showing correct ECG electrode placement on the chest for a 7-lead ECG system (STAT Medical Education, 2014).

The ECG electrodes record the activation of different parts of the heart, which represent different patterns within the ECG data, shown as the blue line in figure 7. Software can then extrapolate additional information from the ECG signal, including the current and average heart rate (HR), shown as the black line in figure 8, and the average interval between R-waves (the large peaks in the blue ECG signal) shown as the green line in figure 8 (GLNeuroTechnologies, 2014).

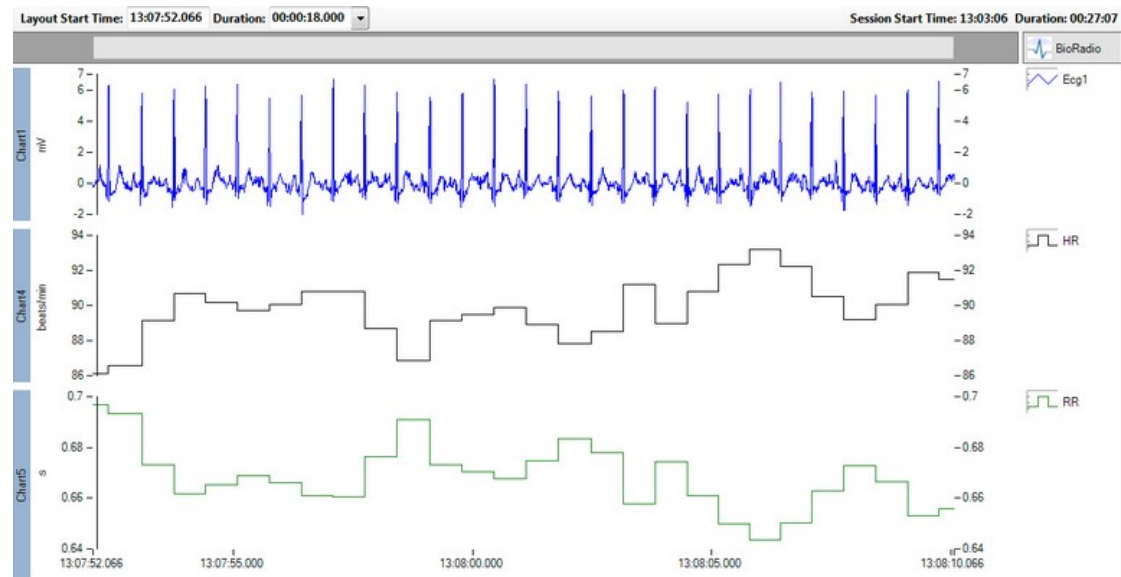


Figure 7: Raw ECG signals and the HR and RR extrapolated from the ECG signal, as recorded through a BioRadio device and software (GLNeuroTechnologies, 2014).

ECG devices have a number of benefits and limitations which may impact upon their selection for use within this project:

### Benefits:

- **Highly Detailed:** ECG provides a wealth of information on the rhythm and status of the heart, including specific influences and the activation of different components of the heart (Cacioppo, Tassinary and Berntson, 2007).

### Limitations:

- **Complicated:** ECG requires precise placement of between 3 and 12 electrodes on the chest. To gain the full benefit of the detail that ECG can provide, at least 6 electrodes need to be used (Cacioppo, Tassinary and Berntson, 2007), although this may be excessive for this project as there is no need for medical-grade detail into heart conditions.
- **Discomfort:** While the electrodes themselves are painless and non-invasive, they need to be placed in precise locations on the chest of the participant. In an experimental setting, the requirement to place the electrodes upon and



immediately surrounding the breast could result in the discomfort of participants.

- **Movement artefacts:** Similarly to EEG, ECG is susceptible to artefacts if participants move excessively. This is due to the chest and arm muscles producing electrical signals when activated, which can confound the measurement of the electrical signals generated by the heart (Cacioppo, Tassinari and Berntson, 2007).
  - This issue is even more pronounced in ECG, given the size of the muscles in the chest and arms, and so the use of ECG would limit the ability for participants to play within a naturalistic environment.

### 2.3.2.2.2. Photoplethysmography (PPG)

Photoplethysmography (PPG) (or Pulse Oximetry) measures the heart rate in a different manner, and involves clipping a sensor onto the tip of a finger. PPG devices direct specific wavelengths of light into the tip of the finger, and then measure the wavelengths and brightness of the light as it exits the finger. Variations in the wavelength and brightness allow the device to determine the heart rate, as well as the oxygen saturation of the blood through measuring the pulse of arterial blood in the finger (Cacioppo, Tassinari and Berntson, 2007).



Figure 8: A portable Pulse Oximeter, illustrating the finger clip and a wrist-device for displaying the heart rate and level of blood.



PPG devices are substantially simpler than ECG devices, and are limited in that they are only able to provide the heart rate and the level of oxygenation, however they have a number of benefits that could make them a more appropriate instrument within video game studies:

### Benefits:

- **Less susceptible to artefacts:** PPG does not measure electrical signals, and so is not susceptible to data artefacts created through extraneous body movement (Cacioppo, Tassinary and Berntson, 2007).
- **Blood Oxygenation:** Blood oxygenation levels collected through PPG can provide additional data on physiological arousal.
- **Comfort:** The finger clip is non-invasive, and does not require electrodes to be placed upon the chest, which might make participants uncomfortable.
- **Simplicity:** The finger clip can be applied within seconds, without requiring precise placement.

### Limitations:

- **No rhythm data:** Due to the way in which PPG devices measure the heart rate, it is unable to provide data on the heart rhythm, nor the influences that are acting upon the heart (Cacioppo, Tassinary and Berntson, 2007).
- **Disruption:** The finger clip is typically placed upon the index finger of the non-dominant hand of the participant in an effort to minimise interference. However, due to the two-handed nature of most video game control devices, the finger clip sensor could interfere with game controls.

### 2.3.2.3. *Skin Conductivity*

Skin is a major body organ that plays a number of roles, amongst which is thermo-regulation, or the regulation of body temperature. One method of thermo-regulation is through perspiration, or releasing moisture onto the surface of the skin in the form of sweat, which is released through sweat glands within the skin (Cacioppo, Tassinary and Berntson, 2007). Since the electrical conductivity of the

skin surface varies depending upon the amount of sweat-induced moisture present, this allows for the precise and objective measurement of sweat gland activation.

The activation of sweat glands is regulated in a similar manner to the heart rate, in that it is influenced both by the parasympathetic and the sympathetic nervous systems (Cacioppo, Tassinari and Berntson, 2007). When the sympathetic nervous system is activated in response to arousing stimuli, one of the physiological responses triggered is the activation of sweat glands within the skin in anticipation of impending physical activity regardless of whether or not the body actually requires cooling at that point in time (Cacioppo, Tassinari and Berntson, 2007). As with heart rate variations, studies have found that video games can result in the activation of sweat glands through the presentation of both threatening and exciting stimuli (Ravaja and Kivikangas, 2008).

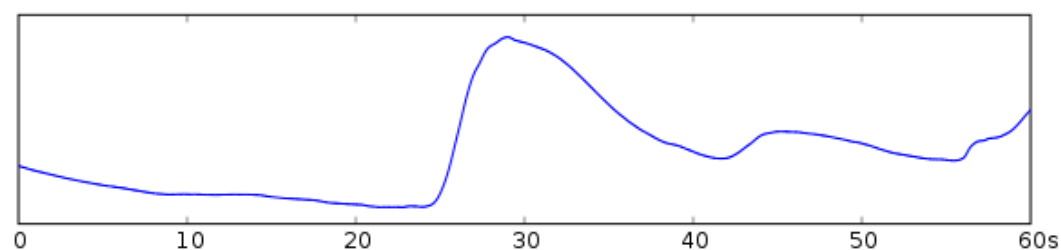


Figure 9: Electrodermal responses to video game stimuli, shown over a 60 second period after the experience of a stimuli (Galvanic Skin Response, n.d.).

Sweat glands are also activated as a result of varying body temperatures, which occurs through the activation of the parasympathetic nervous system to assist in the regulation of body temperature (Cacioppo, Tassinari and Berntson, 2007). This will typically occur when the ambient temperature is causing the body to overheat, or if the physical activities are resulting in the production of excessive heat (Cacioppo, Tassinari and Berntson, 2007). These parasympathetic sweat variations are not of interest for this project, however the influence of the parasympathetic nervous system must be considered within research design, as it could confound the measure of the sympathetic sweat responses that are of interest.

The activation of skin sweat cells is a major indicator of physiological arousal, and can be used as a measure to determine the extent of affective arousal within individuals (Cacioppo, Tassinary and Berntson, 2007). Despite both the heart and sweat glands being jointly influenced by the sympathetic and parasympathetic nervous systems, the varying influences of ambient temperature and physical activity result in independent variations in both sweat gland activation and HR (Cacioppo, Tassinary and Berntson, 2007). The implication is that sweat gland activation and HR do not always directly correlate, irrespective of the fact that they are both indicators of physiological arousal (Cacioppo, Tassinary and Berntson, 2007). This is exacerbated by the delay between the presentations of stimuli and sweat becoming present on the surface of the skin, which can take a number of seconds longer to occur than the heart rate response to the same stimuli, and thereby requires data processing in order to align responses of heart rate with those of skin conductivity (Cacioppo, Tassinary and Berntson, 2007). This illustrates the need to use multiple measures of physiological arousal, so that the extent of sympathetic physiological arousal can be accurately determined.

Studies have successfully shown that skin conductivity does vary while playing video games, especially during experiences that are particularly stressful or exciting (e.g. Hernandez *et al.*, 2014; Koelstra and Muhl, 2012; Drachen, Nacke and Yannakakis, 2010). While these studies have shown that certain aspects of the game play experience can be physiologically arousing in a variety of video games, no studies identified so far have attempted to identify the specific affective responses that are associated with these specific arousing game play experiences. Conversely, a large portion of video game research using skin conductivity focus less on the aspects of game play that influence it, and instead focus on dynamically altering the game experience in response to changes in skin conductivity (Nacke *et al.*, 2011).

### 2.3.2.3.1. Galvanic Skin Response (GSR)

Galvanic Skin Response (GSR) is also commonly referred to as Electrodermal Response (EDR), and involves the use of electrodes to determine the electrical conductivity of the skin surface (Cacioppo, Tassinari and Berntson, 2007). GSR devices use two electrodes, one to apply an extremely low and non-harmful electrical current across the surface of the skin, and the second to measure the difference in current after it has travelled between the two electrodes (Cacioppo, Tassinari and Berntson, 2007).

Where lots of moisture is present, the skin resistance will be lower, and so more current will be received by the second electrode. Conversely, where there is a lack of moisture, the skin resistance will be higher, and so less current will be received by the second electrode. GSR electrodes are often attached to the fingertips of the participant, however they are also able to be attached to alternative locations such as the wrist or feet to limit disruptions caused by having devices in obtrusive locations (Cacioppo, Tassinari and Berntson, 2007).

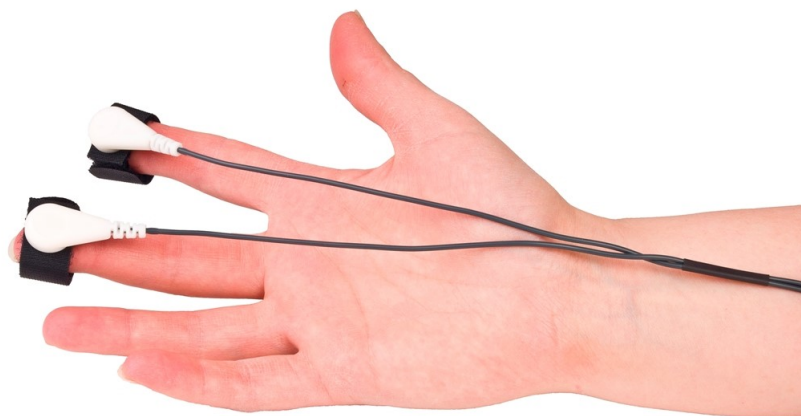


Figure 10: A GSR device attached to the fingers of a participant (YCAN Australia, 2014).

GSR devices have a number of benefits and limitations to be aware of:

### Benefits:

- **Secondary measure:** GSR provides a secondary objective measure of physiological arousal, which will assist with validly assessing the level of affective arousal (Cacioppo, Tassinary and Berntson, 2007).
- **Not susceptible to artefacts:** Although it measures electrical signals, GSR is not susceptible to data artefacts created through movement due to the way in which the electrical signals are measured (Cacioppo, Tassinary and Berntson, 2007).
- **Simplicity:** GSR electrodes are quick and simple to apply, and simply need to be attached in a consistent manner across all participants.

### Limitations:

- **Disruption:** GSR electrodes are often attached upon the tips of fingers on the non-dominant hand. However, due to the two-handed nature of most video game control devices, this could interfere with game controls. Alternative attachment locations can be used to mitigate this limitation.
- **Delayed responses:** GSR signals can take a number of seconds to occur after a stimulus has occurred, which necessitates additional data processing after collection to align physiological responses across the multiple measures used.

#### 2.3.2.4. *Facial Expressions*

Facial expressions are caused by the movement of muscles within the face, and are typically used as non-verbal form of communication to convey emotional states or responses to other people (Cacioppo, Tassinary and Berntson, 2007). Facial expressions can occur both voluntarily at the discretion of the individual, however they can also be produced involuntarily in response to stimuli (Cacioppo, Tassinary and Berntson, 2007).

While it is well recognised that voluntary facial expressions are socially influenced, and can take a wide variety of different forms (Cacioppo, Tassinari and Berntson, 2007), it is currently theorised that involuntary facial expressions may be innate and therefore consistent across all humans (Ekman, 2014). For example, basic facial expressions appear to be consistent across the world, such as those shown in figure 11.

Due to the communicative nature of facial expressions, they have recently seen increasing adoption as a measure of emotional valence in studies utilising psychophysiological instruments. Certain facial expressions are associated with specific affective responses, which allows for facial expressions to be cross-analysed against measures of physiological arousal to determine both the valence and the arousal of affective responses (Ravaja and Kivikangas, 2008). Due to the apparent consistency of basic facial expressions around the world, this provides a largely consistent measure of affective valence.



Figure 11: Different facial expressions, and their associated emotional states (OMRON, 2014).

Psychophysiological instruments largely provide measures of physiological activation, which can then be used to understand psychological effects such as affective arousal. However, without a measure of affective valence there is no simple way for the actual affective response associated with the arousal to be determined (Poels *et al.*, 2012). The measurement of facial expressions can help overcome this limitation, as facial expressions have been shown to correlate closely with affective arousal in repeatedly studies (Cacioppo, Tassinary and Berntson, 2007; Poels *et al.*, 2012). This makes the measurement of facial expressions an important consideration for this project, as it will provide a valid and reliable means of determining the valence associated with measures of arousal.

Studies investigating facial expressions in video games have traditionally used subjective techniques such as video recording the facial expressions of participants and then manually coding these expressions into a number of predefined categories (Cacioppo, Tassinary and Berntson, 2007). While studies using these techniques have had some success in determining emotional responses, they prove problematic for a number of reasons, including inconsistency in coding, observational bias, and the inability to perceive short-lived or subtle expressions that do not produce significant effects upon the appearance of the face (Mandryk and Atkins, 2007).

Given that facial expressions are produced through a series of small muscles within the face, expressions are able to be measured objectively by measuring the electrical activity generated when specific facial muscles activate and analysing these against the facial expressions associated with these muscles. Through objective measurements of facial muscles, the pitfalls of subjective instruments can be avoided, and subtle expressions that might be missed or miscoded by researchers can be more accurately and reliably determined (Cacioppo, Tassinary and Berntson, 2007).

### 2.3.2.4.1. Facial Electromyography (EMG)

Facial Electromyography (EMG) uses a number of electrodes to measure the electrical activity generated by facial muscles as they activate to produce facial expressions, which provides an objective means of determining affective responses that are produced in response to video game components (Cacioppo, Tassinary and Berntson, 2007). The measurement of electrical activity allows for the assessment of facial expressions that are unconsciously produced as part of an affective response, and the precise nature of the EMG device allows for the measurement of affective responses so small that the participant might not be aware that they responded at all (Cacioppo, Tassinary and Berntson, 2007).

EMG electrodes are placed over the location of facial muscles that are most commonly those that are associated with smiling or frowning, with the three most common muscles being the zygomaticus major (controls cheek movements), the corrugator supercilli (controls eyebrow movements), and the orbicularis oculi (controls the eyelid) (Poels *et al.*, 2012; Salminen and Ravaja, 2007). These three muscles are most commonly because they have been highly correlated with emotional valence (Poels *et al.*, 2012), and because they are slightly easier to isolate than other facial muscles, which reduces the measurement of extraneous muscle activations (Cacioppo, Tassinary and Berntson, 2007).



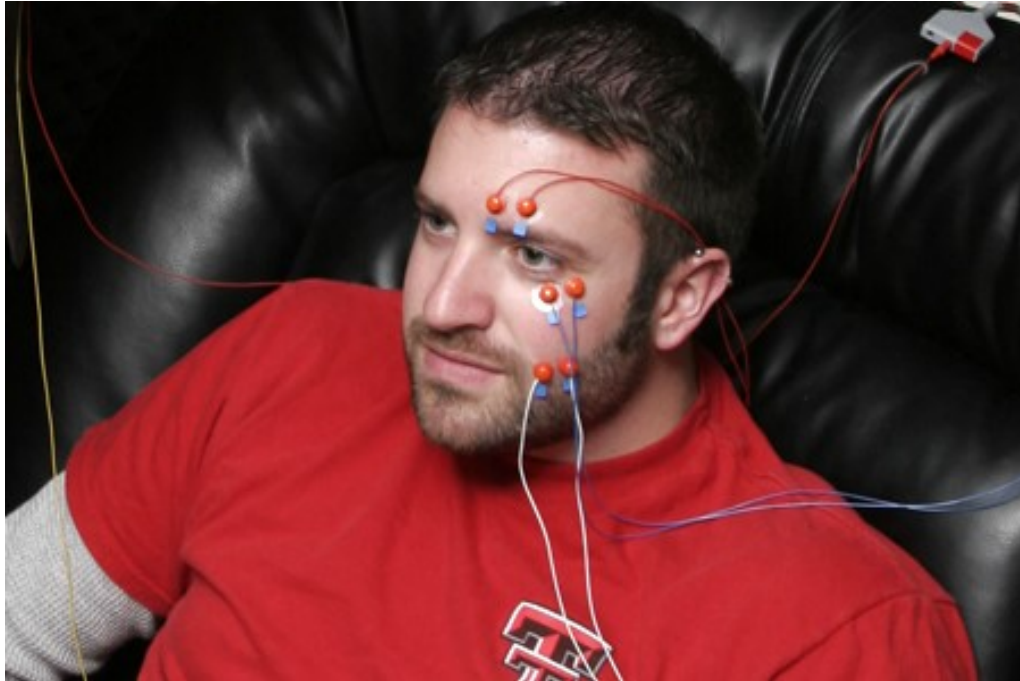


Figure 12: EMG electrodes being used on a participant, showing the placement on the three most commonly measured facial muscles (brow, eyelid, and cheek) (Texas Tech, 2014).

EMG devices have a number of benefits and limitations that need to be considered within the design of research projects:

**Benefits:**

- **Objective measure:** The activation of certain facial muscles has been highly correlated with emotional valence, and so EMG provides an objective measure of valence that avoids many of the pitfalls of subjective measures of valence (Poels *et al.*, 2012).
- **Instantaneous response:** EMG provides data on the instantaneous responses that participants have when experiencing stimuli (Ambinder, 2011). This allows for cross-analysis between specific game events or components and the responses that individuals are showing at those points in time.

**Limitations:**

- **Susceptible to manipulation:** Facial expressions can be produced voluntarily by participants, which can result in manipulated expressions being produced, whether intentionally or otherwise (Larsen *et al.*, 2000).
- **Susceptible to artefacts:** Due to the nature of EMG measurements, excessive extraneous movement can produce artifacts within the data collected, requiring the cleaning and processing of data prior to analysis (Cacioppo, Tassinari and Berntson, 2007).
- **Individual variations:** Although many facial expressions appear to be consistent across populations, many other facial expressions are socially constructed and so vary between individuals (Cacioppo, Tassinari and Berntson, 2007).

### Physiological Responses

Physiological responses refer to aspects of the body that could indicate a physical reaction of some kind when interacting with a video game. These are often measurable using medical equipment such as heart rate monitors and Galvanic Skin Response (GSR). Such physiological responses include:

- Increased heart rate (Wang and Perry, 2006)
- Triggered release of salivary cortisol (Keller *et al.*, 2011)
- Increased sweat activity (galvanic skin response) (Wang and Perry, 2006)
- Increased breathing frequency (Wang and Perry, 2006; Homma and Masaoka, 2008)

All of the above physiological responses are all highly correlated indicators of arousal, which can occur in situations where an individual is active, excited, nervous, or stressed (Kivikangas *et al.*, 2011). The physiological state of arousal (or lack thereof) is an important measure, as it provides researchers with a way of assessing how the body of the individual responds when experiencing or interacting with particular stimuli, which can then provide insights into the how individuals perceive and experience different stimuli.

### 2.3.3. Arousal

Arousal is a multifaceted state that is broadly considered to be a state of physiological alertness, attentiveness, and excitement (Fowles, 1980). Arousal has three primary types, discussed below, which influence the body in differing ways and in response to different stimuli. All three types of arousal are controlled by the reticular activating system (RAC) in the brainstem, and responds to stimuli from both the sympathetic and parasympathetic nervous systems (Cacioppo, Tassinary, and Berntson, 2007).

The implication of arousal states being produced in response to sympathetic nervous responses is that video games could be responsible for producing these arousal states, as there is evidence that video games can trigger such nervous responses (Skosnik *et al.*, 2000). As arousal states result from stimuli responses, while some forms of arousal are relatively universal (such as the fight-or-flight response), others are based upon specific interests of individuals and can vary significantly between people (Greenberg *et al.*, 2010).

#### 2.3.3.1. Physical Arousal

Physiological arousal is a state of bodily activation where the organs are in a heightened state of readiness for physical activity (Fowles, 1980), and includes the release of adrenaline to decrease non-essential bodily functions, increase awareness, decrease sensations of pain, increased heart rate to provide increased oxygen to muscles, and activation of sweat glands to cool the body (Fowles, 1980).

Physical arousal includes deeply rooted unconscious responses to threats, both real or imagined (Jansen *et al.*, 1995), such as the way that people will instinctively jump out of the way of falling objects and will report being in a heightened physical state for some time afterwards. However physical arousal can also be stimulated through positive experiences, such as through cognitive and affective arousal.

Physical arousal works as a function of both parasympathetic and sympathetic nervous system responses, and therefore can occur due to both the fight-and-flight type responses but also due to general bodily functions and homeostatic controls (Jansen *et al.*, 1995). For example, physical arousal can occur due to rising environmental temperature, with the body activating to help disperse and handle the heat before it can cause damage or discomfort.

The factor that renders physical arousal of particular interest to video game studies is that it is far easier to objectively measure physical arousal than it is cognitive or affective arousal, and using relatively simple devices such as heart rate monitors we are able to monitor the extent of physical arousal. Critically, because physical arousal is known to occur as a result of both cognitive and affective arousal (Fowles, 1980), and therefore physical arousal can be used as a proxy measure of these other forms of arousal. The inference of psychological responses based upon physiological responses such as this is a core component of psychophysiology, which is discussed in the following sections.

### **2.3.3.2. Affective Arousal**

Affective arousal is a state of heightened emotional stimulation (Fowles, 1980), such as when individuals are feeling passionate, angry, excited, scared, or happy.

Individuals vary significantly in the extent to which they become aroused affectively, as some individuals can be rather volatile emotionally whilst others can be relatively stable emotionally (Davidson, 1993). Affective arousal is essentially the central spoke of all three arousal states, as both cognitive and physical arousal typically involve some form of emotion during their experience.

Affective arousal is also one of the most interesting arousal states for research involving video games, as there is evidence to show that video games can at least influence, if not directly elicit, emotional responses in players (Ravaja *et al.*, 2009). Alongside cognitive arousal where players are interested and concentrate on the game, affective arousal allows players to experience video games as an emotional

affair, and commonly report feelings such as happiness during or after playing games (Ravaja *et al.*, 2009).

The ability to measure and assess the extent of affective arousal would therefore be a valuable avenue for research, especially in terms of assessing the influence of specific components of video game play on the player emotions.

### 2.3.4. Affective States

Affect can be considered to be the emotions that a person experiences in response to a given set of stimuli (Watson and Clark, 1988), which in this project are the emotions experienced while players interact with specific components of video games. Affective states are by definition short-lived, and will rapidly dissipate after the causal stimuli is removed (Cacioppo, Tassinary and Berntson, 2007; Sims, 1988). In comparison, moods are long lasting dispositions that are shaped over time by a range of influences, including affective states, and which can itself influence the individuals' perception of stimuli (Cacioppo, Tassinary and Berntson, 2007; Sims, 1988).

While assessing the influence of video games on long-term mood is a priority for many studies within this field, these are unable to provide insights into the influence that specific game components have upon this overall mood impact. As this project is proposing to measure the emotional experiences of players as they interact with specific components of video games, such as completing a line or hindering the competing player, it is imperative that the short-term affective responses to video game stimuli are measured as a priority, as this is the true reflection of what the player experiences within the video game at any given point in time (Mandryk and Atkins, 2007).

In relation to the measurement of arousal and valence, this means that to adequately assess the impact of video game components on emotions and affective responses, the instruments used will need to be capable of assessing the strength of

the arousal, the tone of the valence, and a means for assessing whether these are originating from short-term affective responses, or long-term moods.

### *2.3.4.1. Affective Responses to Stimuli*

The most widely accepted model of emotion within psychophysiological studies is Russell's dimensional theory of valence and arousal (Nacke, Drachen and Gobel, 2010; Mandryk, Atkins and Inkpen, 2006; Lang, 1995), which posits that emotion is dynamic rather than rigid, and works through an interplay within the two-dimensional space of arousal and valence (Russell, 1980). Valence, or tone, is indicative of the positivity or negativity of an emotion, such as the difference between being engaged or being bored (Russell, 1980).

Emotional arousal, or activation, can be considered as the strength of the emotion, is an underlying cause of physiological arousal, as emotional states are often expressed through physiological responses, such as increased heart rate and electrodermal activity (Russell, 1980). As can be seen in Figure 13, emotions can be classified within Russell's two-dimensional space depending on the relative valence and arousal of the emotion. For example, Anger is located in the top left hand corner, as it is both highly arousing, and negatively valenced.

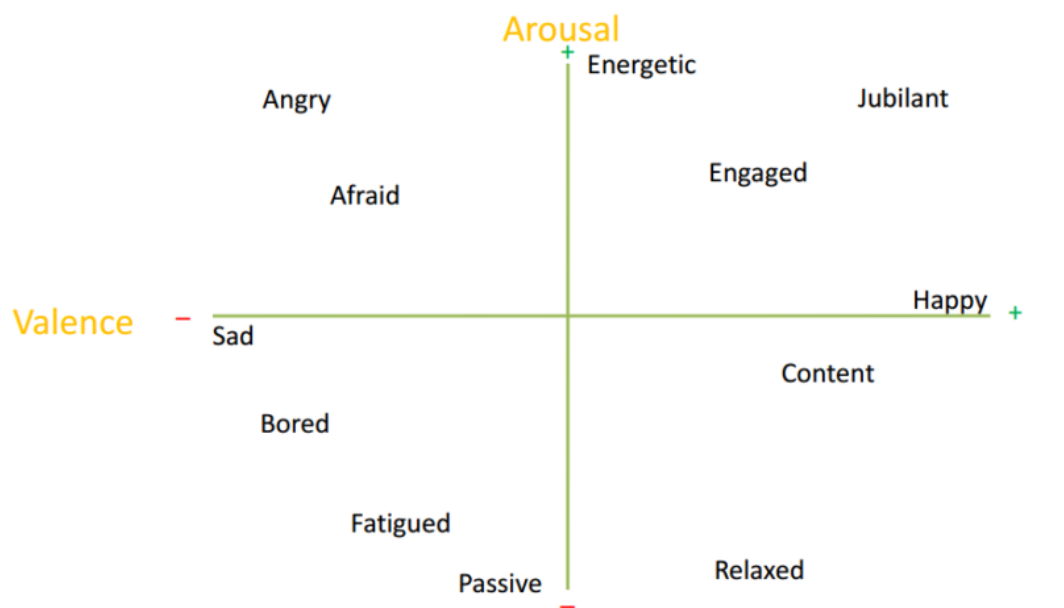


Figure 13: Russell's two-dimensional space with Arousal and Valence, and the emotions associated with different coordinates within the model (Ambinder, 2011).

The popularity of Russell's two-dimensional model of emotion appears to be largely based on the inability of psychophysiological devices to directly measure emotions, which results in a reliance instead on measuring valence and arousal through physiological signals, which can be indicative of emotional states (Cacioppo, Tassinary and Berntson, 2007).

This necessitates a degree of interpretation of the signals in order to identify specific emotions and affective states. For example, an experience that is resulting in high measures of physiological arousal but negative measures of valence could be interpreted as the experience of an intense negative emotion such as anger, frustration, or fear (Ambinder, 2011). To improve the validity and reliability of these interpretations, multiple devices are typically used, including multiple objective devices and signals, and a number of subjective measures such as survey instruments (Cacioppo, Tassinary and Berntson, 2007; Ambinder, 2011).

### 2.3.5. Summary of Psychophysiology

The use of psychophysiological instruments and measures can be shown to have a number of profound benefits in addition to traditional subjective measures, and these can provide this project with significant novel insights into the nature of the affective responses elicited through interaction with varying aspects of video game play. Despite being more expensive and harder to apply than subjective measures such as surveys or interviews, psychophysiological instruments can provide detailed objective data that avoids many of the biases and interpretation issues that subjective instruments can introduce (Kivikangas *et al.*, 2011; Cacioppo, Tassinary and Berntson, 2007). The use of biometric data has the additional benefit of being able to provide data on unconscious and involuntary mental and physiological processes, which it is not possible to capture through subjective instruments (Kivikangas *et al.*, 2011).

Despite the numerous benefits that psychophysiology can provide, there are a number of limitations that need to be factored into the research design. Psychophysiological instruments require a number of electrodes be placed on multiple locations on the body of a participant, which may be distract the participant and result in an atypical play experience for participants (Ambinder, 2011).

Additionally, the equipment used can be cumbersome and difficult for participants to ignore or forget about, which risks distracting participants or at least potentially moderating natural responses. In spite of this apparent issue, no studies were identified that attempted to assess these potential measurement effects within video game domains.

#### 2.3.5.1. General Benefits of Psychophysiology

- **Real-time data collection:** Data can be collected covertly and without interruption throughout the duration of the video game experience



(Kivikangas *et al.*, 2011). This is not possible using subjective measures, where data is typically collected after the experience.

- **High temporal resolution:** Data is collected constantly, with millisecond-level measures providing precise time-lines for varying responses to aspects of video game experiences (Cacioppo, Tassinary and Berntson, 2007). As with the collection of real-time data, this is also not possible using subjective measures.
- **Bias:** Data is collected from physiological signals which, with the exception of facial muscles, are not able to be voluntarily manipulated by participants. This means that psychophysiological data is not subject to response bias, perception bias, or researcher interpretation biases (Kivikangas *et al.*, 2011; Cacioppo, Tassinary and Berntson, 2007).
- **Measures unconscious processes:** Because data is collected from physiological signals, psychophysiological data can provide insights into processes and responses that even the participant is innately unaware of, whether because they were not thinking about it, or because the response was so small that they did not notice (Kivikangas *et al.*, 2011; Cacioppo, Tassinary and Berntson, 2007).

### 2.3.5.2. General Limitations

- **More expensive:** Psychophysiological instruments are more expensive than subjective instruments that provide similar insights. This is due to the requirement for sophisticated and precise devices that are capable of accurately and reliably measuring physiological signals.
- **Time consuming:** Due to the number of electrodes that need to be placed upon the body, studies utilising psychophysiology require additional time to setup equipment for each participant. This could take between 10 and 45 additional minutes, depending upon the type and number of physiological signals being used (Kivikangas *et al.*, 2011; Cacioppo, Tassinary and Berntson, 2007).

- **Numerous electrodes:** When using multiple physiological signals, there could be up to 30 electrodes and associated cables attached to the participant for the duration of the experimental session. These could potentially alter the game play experience for participants, as they may play differently knowing that aspects of their body are being recorded.
- **Data processing:** Psychophysiological data needs to be cleaned and processed prior to analysis. This includes the removal of extraneous artefacts within the data, as well as normalising psychophysiological data against baselines to show the actual impact that experiences or stimuli actually have, which can be time consuming (Cacioppo, Tassinary and Berntson, 2007). Additionally, the physiological data must be mapped to psychological constructs, which requires significant additional effort and interpretation (Mandryk & Nacke, 2016).

### 2.4. Junction of Domains: Video Game Research Using Psychophysiology

Studies in the field of video game research first adopted psychophysiology in 2006, with studies by Mandryk (2006; 2007), Nacke (2008; 2009; 2010), and Ravaja (2006; 2008) pioneering the field. Since then, psychophysiology has become increasingly prolific within the study of video games, with a broad range of studies investigating varying phenomena within video games, including studies finding that video games can elicit a number of both affective, or emotional, and physiological responses within players (Cole, Yoo and Knutson, 2012). These responses are complex and overlapping, and can be indicative of many things, including increased physiological arousal (Nacke, 2015), positively valenced emotions such as joy (Mellecker, Lyons and Baranowski, 2013), and negatively valenced emotions such as anger, addiction, and violence (Sim *et al.*, 2012).

An excessive portion of studies investigating the physiological influences that video games have upon players have been attempting to investigate video game addiction and the impact of violent video games upon players (see Anderson *et al.* 2010 for a meta-analytic review). While these are clearly highly important endeavours, the

inordinate focus upon these areas has resulted in few studies regarding other influences that video games could have upon individuals, including the stimulation of positive emotions, improvements to motivation, encouraging social interactions, improving cognitive problem solving skills, reduction in cognitive stress, providing a highly engaging form of entertainment, and offering a temporary respite from reality.

In addition, while existing studies using psychophysiology have demonstrated that there are a broad range of influences that video games appear to have upon human physiology (Kivikangas *et al.*, 2011; Mellecker, Lyons and Baranowski, 2013), these studies have thus far generally not sufficiently distinguished between the specific aspects or components of video games that appear responsible for these impacts. Studies on violence in video games are the exception to this rule, as these tend to suggest that it is the violent aspects of game play are responsible for later observed anti-social behaviours (Anderson *et al.*, 2010), however these studies do not attempt to identify the influence of events or other aspects within the video games, which may include influential or confounding factors occurring within the game experience.

### 2.4.1. Event-Related Responses

Measuring the specific impacts of individual video game components, or the events occurring during game play, can be problematic because of the way in which video games are designed, as they inherently combine multiple game components within the one event or experience. This can complicate attempts to isolate game components for individual examination, as the removal of some game components fundamentally alters the game play experience, and hence invalidate the measures.

Studies that have attempted to explore the influences of game components have suffered from limited success because of many methodological issues, including the use of subjective measures which may inhibit their ability to determine affective valence and arousal and also typically only captures emotional state at the end of play rather than during the experiences (e.g. Ballard, Visser and Jocoy, 2012).

Additionally, research focusing on specific outcomes such as aggression (e.g. Ballard, Visser and Jocoy, 2012) or flow (e.g. Klasen *et al.*, 2012), while valuable, does not necessarily provide insights into more general video game experiences.

While previous studies using psychophysiology have provided a number of interesting and significant insights into both the emotional, cognitive, and physiological influences that video games can have on players in general, they have thus far not sufficiently examined how specific aspects or components of video games, either jointly or individually, contribute towards these influences. This appears to be because of the dominant use of condition averaging within existing studies, whereby video games as a general experience are measured as an average.

While this both simplifies analytic processes, and can provide a number of valuable insights into video game experiences (Mandryk, Atkins and Inkpen, 2006; Cacioppo, Tassinary and Berntson, 2007; Nacke *et al.*, 2009), it comes at the expense the temporal details, minutia, and specific events such as social interactions or game mechanics encountered during the video game experience. The true power of psychophysiology stems from its provision of continuous and real-time measures of physiological responses to stimuli (Kivikangas *et al.*, 2011), and can thereby be used to assess the respective affective responses at any point in time during the video game experience.

There are a few notable exceptions that have attempted to explore event-related responses using psychophysiology, namely studies undertaken by Mandryk (2006; 2007), Ravaja (2006; 2008), and Weber (2009).

The study conducted by Mandryk and her associates was a pioneering work within the video game domain, and was primarily concerned with the use of psychophysiology as an instrument for the assessment of emotions related to player experiences rather broadly. Their study did describe some event-related emotional responses, highlighting for example that players experienced increased arousal when

scoring a goal in a video game (Mandryk, Atkins and Inkpen, 2006; Mandryk and Atkins, 2007). Describing event-related responses were not however the intention behind the study conducted by Mandryk, and so while the method and study are highly valuable, they do not contribute significantly towards understanding events occurring during video game contexts.

In the case of the study conducted by Weber and associates (2009), the study was primarily concerned with the effects of video game violence, and undertook an in-depth analysis of player choices and how differing antecedents of choices and events produced unique game play experiences for each player. The results of this study revealed a variety of differences in arousal responses to varying game events as they occurred during video game play, and contribute valuable insights into arousal responses to changing events as they occur during game play. However, due to the objectives of the study, the authors only collected physiological data associated with arousal, and so no assessment of affective responses or overall game play experience was possible.

The results of these previous studies therefore highlight the ability of psychophysiology to elicit valuable insights into real-time events as they occur during video game play, but also that current theoretical understandings of such events are not yet highly developed and require further examination.

### **2.4.2. Measurement Effects of Psychophysiology on Game Experiences**

Measurement effects are a ubiquitous aspect of research across domains, whereby the act of observing or measuring phenomena often influences the properties observed (Kassam and Mendes, 2013). In the case of research into video game experiences and emotions, both subjective and objective instruments of measurement are likely to exert an influence upon the experiences of game play being examined.

Emotions are inherently complex, involving multiple physiological systems that interrelate and exert varying influences upon each other depending upon the emotional state and situational contexts (Scherer, 2005; Cacioppo, Tassinary and Berntson, 2007). A core aspect of emotional experiences, particularly as they relate to video games, is that while individuals may be passively aware of emotions, the act of appraising or reflecting upon the emotional experience disrupts these processes (Lambie and Marcel, 2002; Kassam and Mendes, 2013). Additionally, emotional experiences are highly subjective in nature, and require the individual to interpret them, which may differ substantially between individuals (Robinson and Clore, 2002; Cacioppo, Tassinary and Berntson, 2007; Nacke, Stellmach and Lindley, 2011).

There is a wide body of evidence that reflecting upon emotional experiences has a substantial impact upon the current emotional states of individuals (Frattaroli, 2006; Gross, 2002), and the ability to do so serves as the basis for many psychological therapy techniques (Kassam and Mendes, 2013). These influences occur even with relatively simple reflections upon emotions, such as simply being made aware that images will be about positive or negative emotions causing a change in participant ratings of images illustrating varying emotions (Jakymyn and Harris, 2012), or being asked to rate an image as either happy or sad influencing subsequent neural activities (Lieberman *et al.*, 2007).

An examination of current literature reveals that the most common instruments used to measure emotions and experiences within the domain of video game research are surveys requiring reflection upon emotions or experiences. However, given the above discussion it is expected that the use of these instruments exerts a substantial influence upon the reported outcomes. These influences of subjective instruments form a significant portion of the value behind psychophysiological assessment of experiences and emotions, as psychophysiology does not require the participant to be conscious of or reflect upon their emotions (Cacioppo, Tassinary and Berntson, 2007; Drachen, L. E. Nacke, *et al.*, 2010; Kivikangas *et al.*, 2011).

However, while the use of psychophysiology and associated instruments for the measurement may be discrete and covert in terms of recording data relating to the emotions and experiences of individuals (Cacioppo, Tassinary and Berntson, 2007; Ravaja and Kivikangas, 2008a), the requirement to place electrodes on the body of individuals is likely to also influence the experience of participants. Within the context of research into video games, this likely further compounds existing differences between natural video game play and video game play as conducted within controlled experimental settings.

Experimental settings typically require participants to play a specific game (rather than opting to play of their own volition) within an environment that is different to where they might normally play (e.g. in the comfort of their home), and where they are aware of measurements or observations being taken (rather than having privacy).

Within the domain of video game research, some studies using psychophysiology attempt to control against the influence that the equipment may have upon participants, such as Mandryk *et al.* (2007) who attached equipment to both individuals in social game play but only recorded measures from one of the two individuals, partly in an attempt to mitigate the influence that the equipment could have. However, the clear majority of studies ignore the potential measurement effects that the equipment could have upon the game experience, potentially because of the specific measurement effects not being known, or perhaps because they assume that measurement effects must exist but are unavoidable.

### 2.5. Opportunity statement

An examination of extant literature reveals that whilst video games have a number of profound influences upon both the body and mind of players, relatively little is known about the specific influences that game events and contexts have upon these

responses. While the study of emotions and affective responses has seen increased attention in recent years, the complexity of both emotions and video games has left the continuous real-time assessment of affective responses to video game play as an under examined field. The majority of studies that attempt to identify the affective nature of social game play have seen limited success because of their reliance upon limited and often unreliable subjective instruments, as well as being limited to examining only general video game experiences at the completion of game play rather than responses to the nuances of the gameplay experience that occur dynamically during video game play.

Current approaches to objectively measuring affective responses using psychophysiology have shown increasingly promising results, and have greatly expanded understandings of video game experiences and how emotions relate to these. However, due to a range of limitations, these are not sufficient for understanding affective responses within their contextual environments. These limitations include a tendency to average affective responses across the entire gameplay session and hence losing nuances of the experience, attempt to isolate the stimuli of interest and hence lose the context of the experience, or capture continuous data but do not capture sufficient additional contextual data to develop in-depth insights into the conditions within which affective responses are elicited (Nacke, 2013).

The under-examination of affective responses to video game events and contexts within continuous dynamic environments presents a valuable opportunity whereby psychophysiological instruments can be applied in a novel, context-rich, method with the objective of exploring the influence that events occurring during video game play have upon affective responses, as well as overall emotional experiences. A novel method capable of capturing context-rich affective responses to stimuli has the potential to avoid the limitations of both low and high-temporal resolution objective measures. This opportunity extends to research and practice domains beyond those involving video games, whereby the same novel method could be used



to assess affective responses to information systems, design evaluation, and workplace environments, as a few examples.

Developing improved understandings of the affective nature of social video games is of high value both academically and practically. From an academic perspective, it is not possible to fully understand the influence that video games have upon cognition or affect without first developing an understanding of the relationship between affective responses and the specific video game components, contexts, or events that elicit them. While the insights developed throughout this research project are able to contribute to furthering knowledge in this domain, the true academic value in this project is through the development and testing of a novel context-enriched high-temporal resolution approach to measuring and analysing affective responses experienced during video game play, which has the potential to significantly improve the ability of researchers and practitioners to assess responses to stimuli within video games, as well as a range of other dynamic experiential contexts.

The same holds true from a practical perspective, where these same insights are required to assist in the reliable and appropriate design for specific emotional experiences, or to avoid and mitigate specific undesirable responses. This value extends into peripheral fields of video game design and research, including the design of games for serious purposes such as education and health care, as well as the investigation of the impacts that the rapidly growing ubiquity of video game play has upon individuals, culture, and society more broadly.

## Chapter 3: Methodology

### 3.1. Chapter Introduction

This chapter introduces the methodological foundations underlying this research, and identifies the varying methods, techniques, and instruments used in collecting the research data for the purposes of addressing the research questions of this project. The following sections therefore provide a description of the research strategies used to develop the research questions posed within this project, before discussing the selection of instruments and measures capable of addressing them. The pilot studies undertaken are presented, with their results informing the design of the primary experiment of this project.

### 3.2. Introduction

To appropriately address the research objectives of this project, it is necessary to select a research strategy that can enable the researcher to explore pertinent issues in-depth, and within their relative context. To ensure that the data and derived findings are valuable, the research strategies used are required to be valid, rigorous, and appropriate.

Per Crossan (2003: 48), *“the consistency between the aim of the research study, the research questions, the chosen methods, and the personal philosophy of the researcher is the essential underpinning and rationale for any research project.”* To produce consistent and valid approaches to research, it is therefore necessary to align the methods and techniques used to both the specific contexts of the research questions, as well as the philosophical assumptions of the researcher (Proctor, 1998).

Based upon this theoretical foundation of research, the following sections will discuss the underlying philosophical assumptions of the researcher and topic area, leading to the selection of appropriate research methods, instruments, and techniques that align with this philosophical foundation.

### 3.2.1. Research Paradigm

It is widely accepted among researchers that there are a variety of different philosophical positions and assumptions from which researchers approach their research, which are based upon their respective ontological, epistemological, and methodological assumptions (Vasilachis de Gialdino, 2011). Explicating the underlying philosophical assumptions of a research project is of critical importance to the interpretation of intention, as well as the appropriate design and conduct of social science research projects (Walsham, 2006), as they serve to inform the researcher on (Denzin and Lincoln, 1998; Burrell and Morgan, 1979):

- The nature of the world and human reality (*ontology*);
- The type and extent of knowledge that can be gathered from reality (*epistemology*); and
- The ways in which this knowledge can be acquired and understood (*methodology*).

The following sections will discuss the philosophical foundations and assumptions as they relate to this research project.

#### 3.2.1.1. *Ontology of this Project*

Ontology refers to the nature of how humans perceive the existence of the world, and the nature of reality within it (Denzin and Lincoln, 1998; Guba, 1990). While there are dozens of ontologies used across scientific disciplines, most are based upon one of two overarching ontological positions (Crossan, 2003; Orlikowski and Baroudi, 1991):

- *Objective*: The world, and the phenomena therein, exist independently of humans, their actions, and their perceptions; and
- *Subjective*: The world exists within human perception, and can only be understood through the actions and interactions of humans.

These traditionally dichotomous ontological perspectives require the researcher to assess not only their own underlying biases and assumptions, but also whether the nature of the research topic under investigation aims to assess data that can exist independently of human experiences and influences, and should therefore be examined through an objective perspective; or exists as an artefact of human cognition, actions, or interactions, and should therefore be examined through a subjective perspective (Vasilachis de Gialdino, 2011; Tuli, 2010).

However, many phenomena are argued to be far too complex to understand through the dichotomous application of objective or subjective ontological assumptions (Johnson and Onwuegbuzie, 2004; Tashakkori and Teddli, 2010; Zammitto *et al.*, 2014). This is particularly pertinent for research into the human cognition and emotions, which produce physical responses in the body that can be understood and measured external to human experiences, but are simultaneously influenced, experienced, and perceived by humans (Cacioppo, Tassinary, and Berntson, 2007). Per Miller (1978, p. 14), *“the body is the medium of experience and the instrument of action. Through its actions, we shape and organise our experiences and distinguish our perceptions of the outside world from sensations that arise within the body itself.”*

### 3.2.1.1.1. Ontological bases of Psychophysiology

Within fields of research seeking to understand the nature of the mind, including mental functions, properties, and events, and how these relate to the physical body, Philosophy of Mind has historically been the preeminent ontological foundation (Cacioppo, Tassinary, and Berntson, 2007; Braddon-Mitchell and Jackson, 2006). The primary philosophical problem within Philosophy of Mind is the challenge of how the human mind, which is supposedly subjective in nature, can influence the human body, which is objective in nature (Cacioppo, Tassinary, and Berntson, 2007; Braddon-Mitchell and Jackson, 2006). This led to two overarching schools of thought within Philosophy of Mind researchers:

- *Dualism*: Mind and body are ontologically separate entities. Phenomena of the mind are subjective and non-physical in nature, and are irreconcilably different from phenomena of the body, which are objective and physical in nature (Gorham, 1994; Cacioppo, Tassinari and Berntson, 2007).
- *Monism*: Mind and body are ontologically similar entities. Phenomena of the mind are manifestations of the physical brain, and both influence and are influenced by phenomena of the body, which is objective in nature (Gorham, 1994; Cacioppo, Tassinari and Berntson, 2007).

Dualism and its associated schools of thought are no longer widely used within research, with monist schools of thought becoming dominant throughout the twentieth century (Gorham, 1994). The development of modern neuroscience and medical imaging technologies led to a rapid expansion in research efforts investigating the body, with emphasis on the brain and central nervous system, as they pertain to human behaviours and cognition (Cacioppo, Tassinari, and Berntson, 2007). Within the now-dominant monist schools of thought, psychological phenomena and states are considered to have a reciprocal relationship with multiple brain regions, and with additional influences upon the autonomic nervous system, as well as several other body organs (Cacioppo, Tassinari, and Berntson, 2007).

The development of these understandings led a move from physiological philosophical foundations towards the fledgling field of psychophysiology, which added complexity to a monist foundation through the addition of social and cultural influences upon physiological phenomena (Cacioppo, Tassinari, and Berntson, 2007). Within the ontological foundation adopted by researchers within the field of psychophysiology, cognition, behaviours, and emotions are now broadly recognised as being informed or even determined by wider social, cultural, and interpersonal contexts (Cacioppo, Tassinari, and Berntson, 2007).

Psychophysiology, which this project adopts, assumes a unique ontological foundation within the broader monistic Philosophy of Mind school of thought, positing that the human body exists within an objective reality, but also that the body and its physiological responses is inextricably interrelated with subjective human factors such as social and cultural interactions (Cacioppo, Tassinary, and Berntson, 2007). In summary, the ontological foundation of psychophysiology assumes that the body is objective, that the mind is inseparable from it, and that the mind both influences and is influenced by the body. It also posits that the mind, and therefore the body, is influenced by subjective factors and events including social interactions and personal contexts. For the purposes of this research project, this ontological approach suggests that valuable insights into human experiences of the mind could be acquired through the measurement of physical human experiences and bodily responses.

### 3.2.1.2. *Epistemology*

Epistemology refers to the assumed relationship between the phenomena of interest and what can be known about it. In the context of research, this is therefore the philosophical assumption of how valid knowledge on any given phenomenon can be acquired, assessed, and understood (Orlikowski and Baroudi, 1991). Epistemology therefore builds upon an ontological philosophical foundation, and together they serve as a paradigmatic basis for the design of an appropriate methodological approach (Vasilachis de Gialdino, 2011).

There are two overarching epistemological paradigms that are traditionally considered and approached dichotomously, although there are dozens of derivative research paradigms that draw inspiration from one, the other, or a mixture of both (Myers, 1997). The two overarching paradigms are:

- *Positivism*: Phenomena are understood through empirical measurements, independent of the researcher or instruments; and

- *Interpretivism*: Phenomena are understood through examination of social constructs such as shared meanings, languages, behaviours, dependent upon the interpretations of the researcher.

The positivist epistemological position is based upon a long tradition of research in the natural sciences, where it has been used as the philosophical bases of research that tests hypotheses through empirical measurements and experimentation (Orlikowski and Baroudi, 1991). According to Myers (1997), positivist research tends to test theories in the attempt to provide predictive understandings of phenomena.

The interpretive epistemological position is a more recent philosophical position within research contexts (Walsham, 2006). Interpretive research has been developed as an alternative to positivist research, which is criticised as being inappropriate for providing insights and understandings into phenomena that are social in nature (Mingers, 2001). Rather than attempting to measure and describe phenomena empirically, interpretive research instead attempts to develop a rich understanding of it through interpretation of social constructs (Kaplan and Duchon, 1988).

As with ontological foundations, the complexity of research involving humans is inherently complicated, with many interrelated and conflicting factors. When considering research objectives such as developing understandings of emotions within social contexts, the overarching and typically dichotomous objective and subjective paradigms do not afford the flexibility or holistic ability to develop understandings as they relate to the human condition (Johnson and Onwuegbuzie, 2004; Cacioppo, Tassinary and Berntson, 2007; Zammitto *et al.*, 2014). While the body may exist physically, and respond in ways that can be measured empirically, physiology is known to respond to subjective factors, such as social interactions and emotions (Cacioppo, Tassinary and Berntson, 2007; Kassam and Mendes, 2013).

### 3.2.1.2.1. Epistemological basis of Psychophysiology

The epistemological assumptions of psychophysiology build upon its previously discussed monistic ontological foundation, which assumes that the body exists within an objective reality, but also that the body and its physiological responses is inextricably interrelated with subjective human factors such as social and cultural interactions.

Psychophysiology contains several epistemological positions, which can essentially be condensed into reductive or non-reductive positions, both of which assume that mental conditions supervene on physiological states, and that there can be no change in mental states without changes in the physical state (Cacioppo, Tassinari and Berntson, 2007):

*Reductive:* Holds that all mental conditions and properties can be explained through scientific examination of physiology, and its respective responses and processes.

*Non-Reductive:* Holds that mental conditions cannot be reducible to physiological conditions, due to differences in languages and theories used to understand each.

The field of psychophysiology now predominately holds a reductive epistemological position, with researchers working to disentangle and explicate the relationships between varying psychological and physiological processes (Cacioppo, Tassinari and Berntson, 2007). While psychophysiology is predominately positivistic, due to the holistic influences that act upon the human body and all aspects of the human condition, it necessarily draws additional insights from the interpretivist paradigm. It therefore uses objective measures to improve knowledge on physiological processes and mental phenomenon associated with them, however also often uses subjective instruments to contribute improved depth of knowledge regarding how these mental and physical processes are perceived, experienced, and influenced by individuals (Cacioppo, Tassinari and Berntson, 2007).



In summary, the epistemological paradigm of psychophysiology assumes that psychological phenomena can be measured and analysed objectively through disentangling and explicating the relationships between physiological states and responses. It also posits that physiology is influenced by social and cultural interactions and processes, which can best be understood through subjective measures and analysis, creating a holistic approach to understanding the human condition within its contextual environment.

### 3.2.2. Research Objectives

This research project aims to contribute new understandings into the use of psychophysiology within the domain of video game research, with a focus on furthering knowledge on the influence that individual stimuli such as video game events can have upon affective responses.

Existing studies within the video game domain predominately use subjective instruments, with a host of associated issues previously discussed such as response biases and predominately capturing the overall perceived experience rather than specific events or contributing factors that might nuance the experience during game play. Of the studies within the video game domain using psychophysiology, predominately low-temporal resolution methods are used, where the game experience and affective responses are typically averaged per gameplay session (e.g. (Ravaja *et al.*, 2005; Nacke and Lindley, 2009; Drachen, L. E. Nacke, *et al.*, 2010; Klarkowski *et al.*, 2016)). While this does capture data on the specific events and contributing factors that nuance the game experience, it does not provide any avenue for exploring these nuancing factors, as they are averaged out across the play session. Some attempts have been made to use high-temporal resolution methods, such as Mandryk *et al.* (2006;2007) who used a host of psychophysiological measures to assess and model player emotions, with the ability to provide insights into emotions at any point during video game play.

However, Mandryk *et al.* (2006) conducted their study with the intention of developing and validating a model of emotions for use in future research, and did not specifically assess event related affective responses. Additionally, Mandryk *et al.* (2006) assessed affective responses by modelling physiological measures of arousal and valence into a series of 5 categories of emotions, but did not attempt to add further contextual information that could provide insights into game events or contextual environments. Additional contextual information of value could include social interactions between players, analysis of gameplay events, or the influence that the presence of psychophysiological instruments and electrodes had upon participant responses.

To this end, this project required the use of a novel context-enriched high-temporal resolution method for the analysis of individual video game stimuli and their respective affective responses within a naturalistic contextual environment, without requiring the isolation of specific stimuli. This was intended to allow for the capture and analysis of a holistic set of events and nuancing factors that could contribute towards the overall game experience, including affective responses to each stimuli of interest.

### 3.2.3. Research Questions

#### *Exploratory Research Problem (RP):*

*How can affective responses to real-time video game events be objectively measured using a novel context-rich psychophysiological method?*

In summary from the literature review presented within Chapter 2, while it has been established that social video game play can produce a range of emotional experiences over the course of the game session, such as increased excitement and enjoyment (Sweetser and Wyeth, 2005), increased frustration (Andrew K. Przybylski, Rigby and Ryan, 2010), and changed self-awareness (Christoph, Dorothée and Peter, 2009), these are typically measured as an end state of the overall video game

experience. While the use of real-time objective measures has seen increasing popularity (Kivikangas *et al.*, 2011; Cowley *et al.*, 2016), the complexity of emotions and video games has resulted in relatively few studies examining video game events and their respective influences upon affective responses, and there is certainly much more to be learned within this space.

Several studies have used psychophysiological instruments with the aim of recording the affective states of players within video game contexts, with evidence that such instruments can reliably be used as a measure of overall player experience and to assess levels of fun and enjoyment (Ravaja *et al.*, 2005; Mandryk and Atkins, 2007; Salminen and Ravaja, 2007; Drachen, L. E. Nacke, *et al.*, 2010; Nacke, Stellmach and Lindley, 2011; Ben Cowley *et al.*, 2014). However, current studies have either not attempted or insufficiently explored the dynamic affective responses to game events as they occur in real-time during video game play. It is therefore not well understood how factors such as game events occurring during video game play, social interactions, or even the presence of physiological measurement instruments contribute towards the overall composite video game experience.

The events that this study intends to examine for the respective influence that they have upon affective responses are those that could be commonly expected to occur within naturalistic (a typical for social video game play) contexts during video game play. These include interactions between players, whether it be verbal or physical, and events that occur during the game play, including differentiating between events that only directly influence one or the other player, but which might exert indirect influences upon the player that did not experience the event. Worth considering also are false positives, whereby players may recall pleasant experiences external to the game, which may influence measures of affect or experience.

Additionally, while the use of psychophysiological instruments has seen increasing use within the domain of video game research, no existing studies have been identified that attempted to assess the influence that such instruments have upon

the experience of players during video game play. Given that many studies using psychophysiology require placing multiple electrodes upon the face of participants, multiple electrodes upon the hands of participants, and potentially several electrodes upon the chest of participants, it is likely that participant awareness of the experimental setting would be increased. It is therefore plausible that this increased awareness may exert an influence upon the affective responses experienced during video game play, potentially reducing the reliability of methods using psychophysiology.

To address this research problem, the following exploratory research questions need to be addressed:

### ***Research Question 1 (RQ1):***

*How do events occurring during video game play influence event-related affective responses?*

Whilst previous research has shown that video games in general can produce a range of affective responses within players, it is currently not understood how events experienced during video game play contribute towards these responses, nor to what extent these events influence the overall affective experience. This warrants an examination of events experienced during video game play using a range of both objective physiological and subjective additional measures, assessing events as the unit of analysis, facilitating the examination of dynamic real-time affective responses elicited through specific events experienced during game play.

### ***Research Question 2 (RQ2):***

*How does use of psychophysiological instruments influence the player experience during video game play?*

The use of psychophysiological instruments is increasingly common within the field of video game research, particularly for the investigation of player experiences and general emotional experiences. While it is typical for studies using physiological measures to use questionnaires assessing perceived intrusiveness of measures (Nacke, 2016), it stands to reason that the use of physiological measures may influence play experiences due to changes in self-awareness or perceptions due to participants feeling observed. However, no studies identified have sought to identify the influence that the use of such instruments might have upon the player experience, whether by increasing player awareness of the study context, or by causing disruptions within the play environment. It is therefore not currently known how the use of psychophysiological instruments alters the perception of players and therefore the game play experience, nor whether this translates to altered affective responses. It is understood from studies within other fields that heightened participant awareness of the study, its equipment, and the feeling of being observed all contribute towards changed behaviours (Lieberman *et al.*, 2007; Kassam and Mendes, 2013), so it is plausible that these effects carry over to psychophysiological studies, however this has not previously been tested empirically within the domain of video games research.

### 3.2.4. Data Requirements

Adequately addressing the research problem and its respective research questions requires data from a variety of sources, each with specialised data processing requirements. An overview of the research questions and their respective requirements are illustrated in Table 1 below:

Research Question:	Information Required:	Physiological Data:	Additional Data:	Derived Measures:
<b>1: How do events occurring during video game play influence event-</b>	Real-time measurement of Affective Responses;	EMG <sub>smile</sub> ; EMG <sub>frown</sub> ; GSR; HR.	Facial Observations; Social Observations;	Derived Affective Responses @

<i>related affective responses?</i>	Temporal recording of Events experienced.		Video Game Events; GEQ; SAM.	Observed Events.
<b>2: Does the use of psychophysiological instruments influence the player experience during video game play?</b>	Subjective measurement of Affective Experiences;  Subjective measurement of Player Experiences.	Recorded, but not required.	Facial Observations; Social Observations; Video Game Events; GEQ; SAM; Demographics.	Player Experience; Subjective Affect @ Psychological Presence.

*Table 1: Summary of Research Questions and their respective data requirements, explained further below.*

To appropriately address both of the aforementioned Research Questions, a total of 9 data sources were used for collection. Each of these are more thoroughly discussed in later sections of this methodology chapter, however they are briefly outlined below:

- Electromyography Frown (EMG<sub>frown</sub>): Electrodes placed over the *corrugator supercilii* muscle, on the inner eyebrow. Measures negatively valenced affect (Cacioppo, Tassinary and Berntson, 2007; Boxtel, 2010; Cowley *et al.*, 2016).
- Electromyography Smile (EMG<sub>smile</sub>): Electrodes placed over the *zygomaticus major* muscle, on the cheek. Measures positively valenced affect (Cacioppo, Tassinary and Berntson, 2007; Boxtel, 2010; Cowley *et al.*, 2016).
- Galvanic Skin Response (GSR): Electrodes placed over the *thenar* and *hypothenar* muscle sites, on the palm of the players non-preferred hand. Measures physiological arousal, positively correlated as a key indicator of affective arousal (Cacioppo, Tassinary and Berntson, 2007; Cowley *et al.*, 2016).

- Heart Rate (HR): Photoplethysmography (PPG) Probe placed upon the tip of players non-preferred index finger. Measures physiological arousal, positively correlated as a key indicator of affective arousal (Cacioppo, Tassinary and Berntson, 2007; Boxtel, 2010; Cowley *et al.*, 2016).
- Facial Expression Observations (FO): Video recording of player's face made during video game play. Facial expressions codified upon a timeline for comparison against other measures.
- Social Interaction Observations (SO): Video recording of inter-player interactions during video game play. Interactions codified upon a timeline for comparison against other measures.
- Game Event Observations (GEO): Video recording of events and mechanics during video game play. Events codified upon a timeline for comparison against other measures.
- Game Experience Questionnaire (GEQ): Survey of perceived self-reported video game experience, completed at the end of each experimental condition (IJsselstein, de Kort and Poels, 2008). Responses summarised for comparison against perceived affect, experiences, and attitudes towards game play.
- Self-Assessment Manikin (SAM): Survey of perceived self-reported affect, completed at the end of each experimental condition (Bradley and Lang, 1994). Responses summarised for comparison against physiological measures affect.

### **3.2.4.1. Data Requirements: Research Question 1**

Research Question 1 aims to collect data with the intention of analysing the affective responses of players at a temporal level, comparing measured affective responses at any given point in time against events of interest, either within the video game itself (such as players interacting with game components), or within the social environment (such as players interacting with each other).

To this end, Research Question 1 required the recording of data from a total of seven sources: EMG<sub>frown</sub>, EMG<sub>smile</sub>, GSR, HR, FO, GEQ, and SAM. These data sources are used in an amalgam, allowing for the analysis of affective responses derived from a composite of physiological and subjective measures, rather than using raw values which are often not comparable between participants (Cacioppo, Tassinary and Berntson, 2007; Kivikangas *et al.*, 2011; Cowley *et al.*, 2016). The derived affect is analysed against observed video game events (including interaction with game components, and interactions with other players), allowing for comparisons of affective responses against similar or dissimilar observed events.

### **3.2.4.2. Data Requirements: Research Question 2**

Research Question 2 aims to collect data with the intention of analysing the influence that the presence of physiological measurement devices has upon the relative player experience and affective responses during video game play, potentially conveying insights into the efficacy of such instruments and furthering understandings of how appropriate they might be for future video game research in these fields.

To this end, Research Question 2 required the recording of data from a total of eight sources: EMG<sub>frown</sub>, EMG<sub>smile</sub>, GSR, HR, FO, SO, GEQ, and SAM. These data sources can be amalgamated or used individually to provide a range of lenses through which influences that the physiological equipment had upon research participants.

### **3.2.5. Ethics**

Due to the nature of the BioRadio device (discussed in Section 3.3.1) and the selected physiological signals (discussed in Section 3.4), which are lightweight, non-invasive, and uses passive measures, this project does not involve any notable ethical concerns. A minimal risk ethics application was submitted for approval with the University of Tasmania Human Research Ethics Committee, and this approval was granted (approval number H14570) before the commencement of participant recruitment.



To maintain compliance with the ethical requirements for human research, all data was stripped of identifiable information and anonymised through the use of coded pseudonyms, and stored on a secure University server. Information sheets were provided to participants prior to the commencement of experimental sessions informing participants of their rights and details of the research, and consent forms were collected from participants and then stored in a private location on campus.

### 3.3. Instruments

This research project used a range of instruments to facilitate the valid and reliable capture of all data necessary to address its research objectives. These include a range of psychophysiological instruments, a video game, and devices for the video recording of participants and gameplay videos. The following section discusses the selection of these instruments.

#### 3.3.1. Psychophysiological devices

Two psychophysiological devices, the BioRadio by GLNeuroTechnologies and the EPOC+ by Emotiv, were used for the purposes of this research project.

The BioRadio device, displayed in Figure 14, is capable of simultaneously recording a wide range of psychophysiological signals, including brain activity (EEG), heart rate (instantaneous and variability) (HR and HRV), electrodermal activity (GSR), activation of facial muscles (EMG), rate of breathing, body temperature, and body movements. Additionally, the BioRadio device is relatively inexpensive, only slightly larger than a mobile phone, battery powered, and highly portable. This renders the device highly suitable for use within this study, as it allows for the research to be conducted within a more naturalistic game play environment rather than the traditional psychophysiological laboratory setting, which could significantly alter the psychophysiological measures. The BioRadio requires that a series of sensors be placed upon the body of participants in predetermined locations, in order to capture the desired psychophysiological signals. Descriptions of the chosen signals, as well as

the placements of the sensors required to capture them, are discussed in the following sections.



Figure 14: BioRadio device on arm (GLNeuroTechnologies, 2014).

The EPOC+ EEG headset was selected for use to measure EEG signals during a pilot experiment conducted within this project, as the headset *allowed* for significantly faster experiment setup on participants and was more comfortable for participants as it used saline solution on its contacts rather than conductive gel which would require participants to wash their hair afterwards. The EPOC+ device does have several notable disadvantages, in that it only uses 16 channels to capture EEG signals, and that it may introduce additional signal noise due to having less secure electrode contact on the skull, and increased movement artefacts. Since each channel is used for the placement of an electrode, this means that it has a lower spatial resolution for pinpointing the specific origin of brain activity patterns compared to dedicated laboratory EEG devices, which allow between 32 and 256 channels to be used.

Despite this lower resolution, recent studies have found that 16 channel EEG devices are valid instruments that are capable of identifying and categorising brain activity patterns. See Figure 15 below, which compares the classification ability of a laboratory grade 64 channel device against a portable 16 channel device (Duvinage *et al.*, 2013). In all tests, the classification ability between the two devices varied between 1.4% and 3.4%, and both devices were able to classify brain activities with a significant degree of accuracy, which demonstrates that the lower resolution of the EPOC+ device will not have a detrimental influence upon its ability to validly record EEG data.

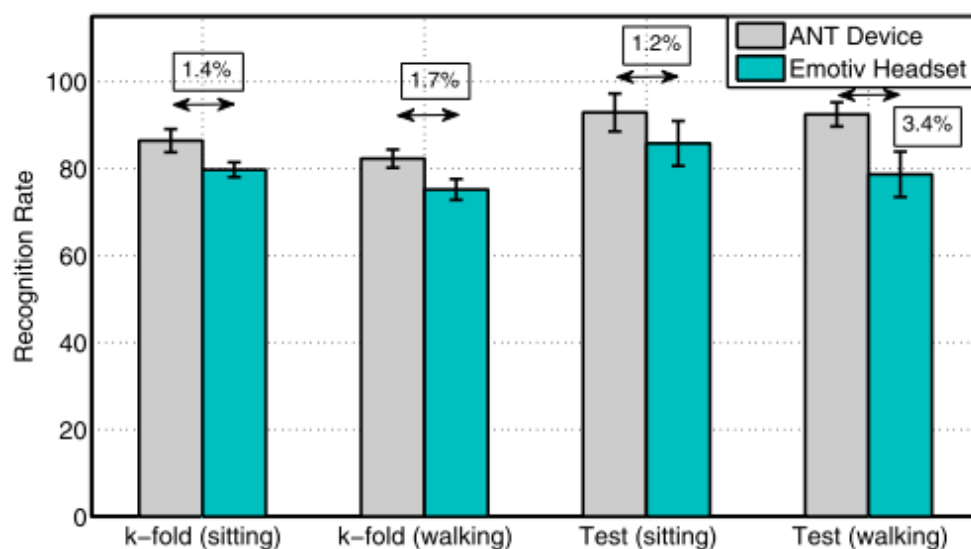


Figure 15: Comparison of 64 channel laboratory EEG device against a 16 channel portable EPOC+ EEG device (Duvinage *et al.*, 2013).

Figure 16 demonstrates how the electrodes for the EPOC+ EEG headset device are placed upon the head of participants. Note that the electrodes are not likely to subject the participant to discomfort, and are therefore unlikely to significantly impact upon the play experience. Compare this with figure 17, which demonstrates the placement of EEG electrodes in traditional psychophysiology devices. In order to validly and reliably record only those psychophysiological signals that pertain to the video game experience rather than those that originate through environmental discomfort or other potential confounds, it is essential that the testing environment

be made as comfortable and naturalistic as possible for the participant (Cacioppo, Tassinary and Berntson, 2007).



Figure 16: EPOC+ EEG headset placed on head of participant (Emotiv, 2015).



Figure 17: Traditional EEG and EMG electrodes placed on head of participant.

It is therefore evident that a 16 channel EEG device is both sufficient and appropriate for the purposes of this research project, when the relative comfort of participants and that the lower spatial resolution produces only slightly lower quality data. This

selection is especially pronounced given the relative difference in price between the two devices. A 64 channel EEG device costs between \$40,000 and \$60,000, requires a dedicated laboratory in which it can be installed, and is only able to measure one type of psychophysiology signal. In comparison, the BioRadio device costs between \$6,000 and \$8,000, is portable, and can also validly measure a wide range of other psychophysiology signals (GLNeuroTechnologies, 2015), and the EPOC+ EEG headset costs between \$1,000 and \$1,200, and is also portable and can validly measure brain activity (Emotiv, 2015; Duvinage *et al.*, 2013).

### 3.3.1.1. Software Used

#### Signal Capture

The BioRadio device was provided with proprietary BioCapture software that allowed the BioRadio to be configured appropriately, and recorded psychophysiological signal data captured by the device (GLNeuroTechnologies, 2015). BioCapture allowed for recorded signal data to be examined, cleaned, managed, and facilitated basic preliminary data checking and processing. BioCapture also facilitated exporting the data in common formats to allow for more complete processing and analysis in external software packages, such as MATLAB, and SPSS.

The EPOC+ device comes with the proprietary Emotiv Xavier software (Emotiv, 2015) that allowed the EPOC+ to record and display EEG signals in real time, and play back previously recorded signals. Xavier allowed for the data to be exported in a range of formats for more complete processing and analysis in external software tools such as MATLAB, SPSS, and EEGLAB.

For further information on the external software packages used to process and analyse the data, please refer to the following section.

### **3.3.1.2. *Signal Processing and Analysis Software***

Due to the nature of the bio-signals used within this research project, a number of software packages are used independently to facilitate the processing and analysis of data from varying sources.

#### **3.3.1.2.1. MATLAB**

MATLAB is a highly extensible software tool that allows for the exploration and visualization of data, and most importantly it allows for software tools to be created, shared, and run to facilitate a wide range of different data processing and analysis needs. A number of open source MATLAB toolboxes are freely available that enable the in-depth processing and analysis of psychophysiological bio signals, and those used are described below.

##### **3.3.1.2.1.1. Toolbox: EEGLAB**

EEGLAB is a popular open source interactive MATLAB toolbox that is used exclusively for the processing and analysis of electrophysiological data. EEGLAB incorporates advanced processing and analysis features such as independent component analysis, time/frequency analysis, artefact rejection, event-related statistics, as well as a range of data visualization modes and tools (EEGLAB, 2015).



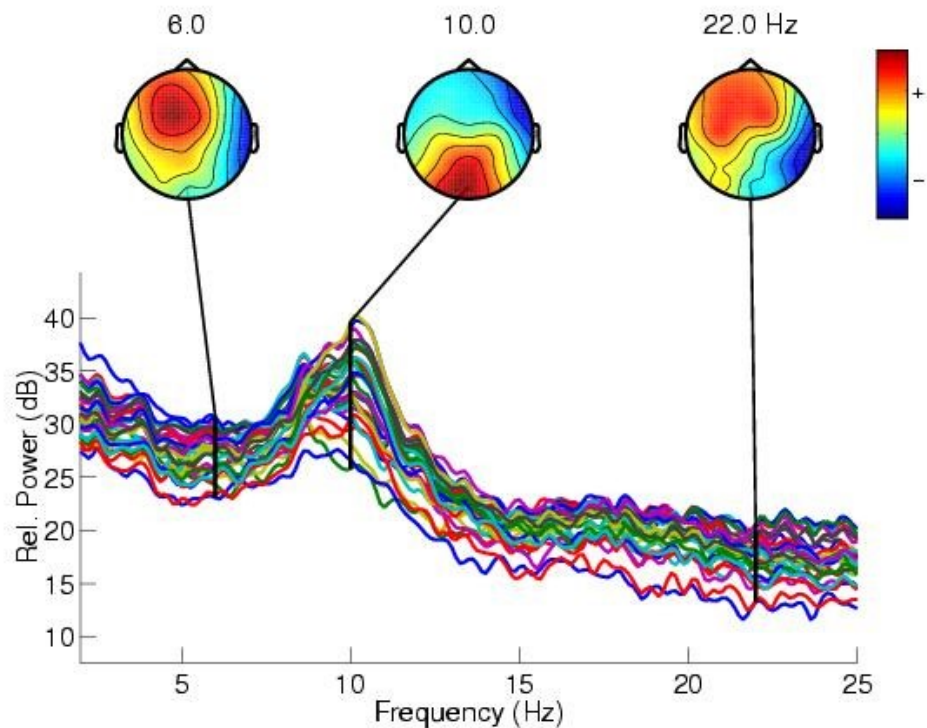


Figure 18: EEGLAB visualization of time/frequency event-related EEG data (EEGLAB, 2015).

A range of other EEG analysis software is available, however these are most often commercial, proprietary, and expensive, such as the popular BESA Research software which is widely used within clinical research settings but costs between \$5,000 and \$15,000 USD depending upon the packages required. Additionally, proprietary software is often not well documented and if support is required it typically costs extra.

In contrast, EEGLAB is free, open source, well documented, and has an active development and research community who are provide support freely. Paid software is typically a little easier to use and requires less training, however EEGLAB has the majority of features that most proprietary EEG analysis software provides and so is considered to be an appropriate choice given the constraints of this research project. EEGLAB was consequently selected for use in this study.

### 3.3.1.2.1.2. Toolbox: LEDALAB

Similarly to EEGLAB, LEDALAB is a popular open source interactive MATLAB toolbox that is used for the processing and analysis of skin conductance data. LEDALAB is capable of importing multiple common file formats, and provides a good range of processing and analysis functions, including decomposition of its tonic and phasic components. This includes the ability to perform Continuous Decomposition Analysis (CDA) and Discrete Decomposition Analysis (DDA), which allows for differential processing and analysis depending upon the dataset and requirements of the research (LEDALAB, 2015).

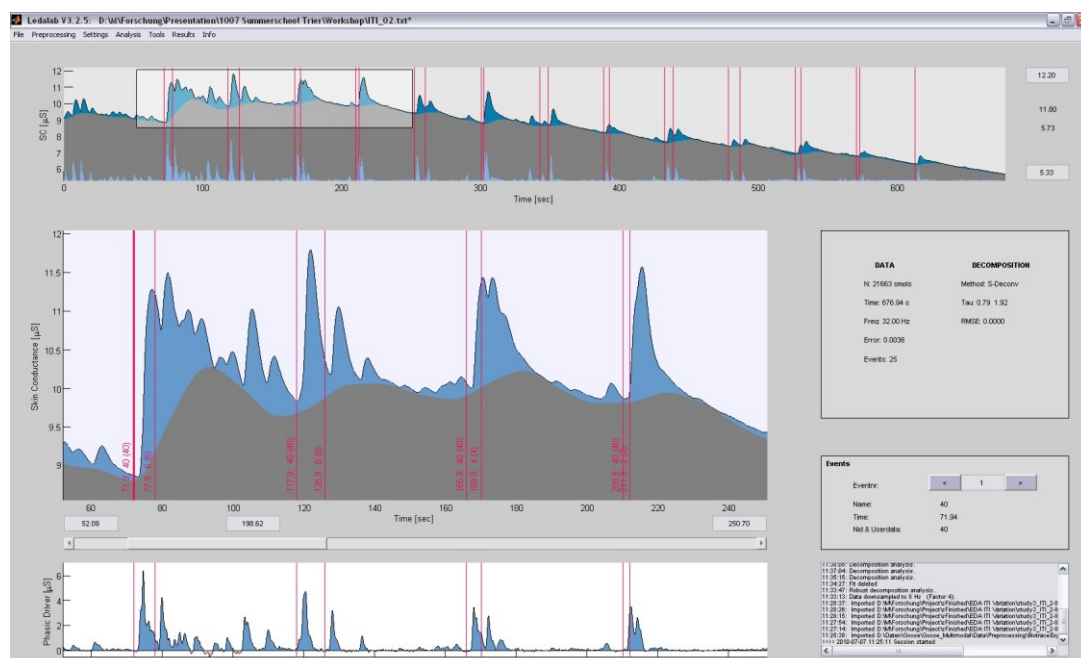


Figure 19: LEDALAB visualization of electrodermal activity over time during event stimuli (LEDALAB, 2015).

There are a range of alternatives to LEDALAB that could be considered, such as the commercially available BIOPAC Electrodermal Activity Analysis software (BIOPAC Systems, 2015). However, much like the commercially available EEG software, it is expensive, proprietary, and not well document. Of the freely available open source software, LEDALAB appears to be the most commonly used and offers a sufficient range of features that make it an appropriate choice for this research project.



### 3.3.1.2.1.3. Toolbox: EMGLAB

Similarly to both EEGLAB and LEDALAB, EMGLAB is a popular open source interactive MATLAB toolbox that is used for the processing, decomposition, and analysis of electromuscular activity. EMGLAB decomposes EMG signals into their component motor-unit potentials (MUP), which facilitates the isolation and analysis of individual muscle activation. This is particularly important when using facial EMG signals as the muscles in the face are small and in close proximity, and therefore require additional care in processing to ensure interfering activities are subtracted from the signals (McGill, Lateva, and Marateb, 2006).



Figure 20: EMGLAB visualization of electromyographic activity over time (EMGLAB, 2015).

As with the other bio-signals, there are a range of commercially available software available that could be used as alternatives, some of which are easier to use and offer some more advanced features, such as EMGworks (Delsys, 2015). However, commercial EMG software all have the same cost and documentation issues as

software packages for other bio-signals, which render them inappropriate for this research project.

The EMGLAB software is one of the most commonly used open source EMG processing and analysis tools, and offers a range of features sufficient for the processing and analysis of EMG data in this research project, which makes it an appropriate choice for this research project.

### 3.3.1.2.1.4. Toolbox: HRVAS

HRVAS is a popular open source interactive MATLAB toolbox that is used for the cleaning, processing, and analysis of heart rate activity across a range of measures and techniques (Ramshur, 2010). HRVAS is capable of importing a range of common heart rate data formats, and includes a range of useful processing and analysis functions, including both average and temporal breakdowns of instantaneous heart rate, average and temporal heart rate variability, and decomposition of heart rate activity across a number of frequency bands to assist in the analysis and interpretation of the data. HRVAS also includes the ability to detect and remove artefacts automatically, based upon either predefined or customised thresholds (Ramshur, 2010).

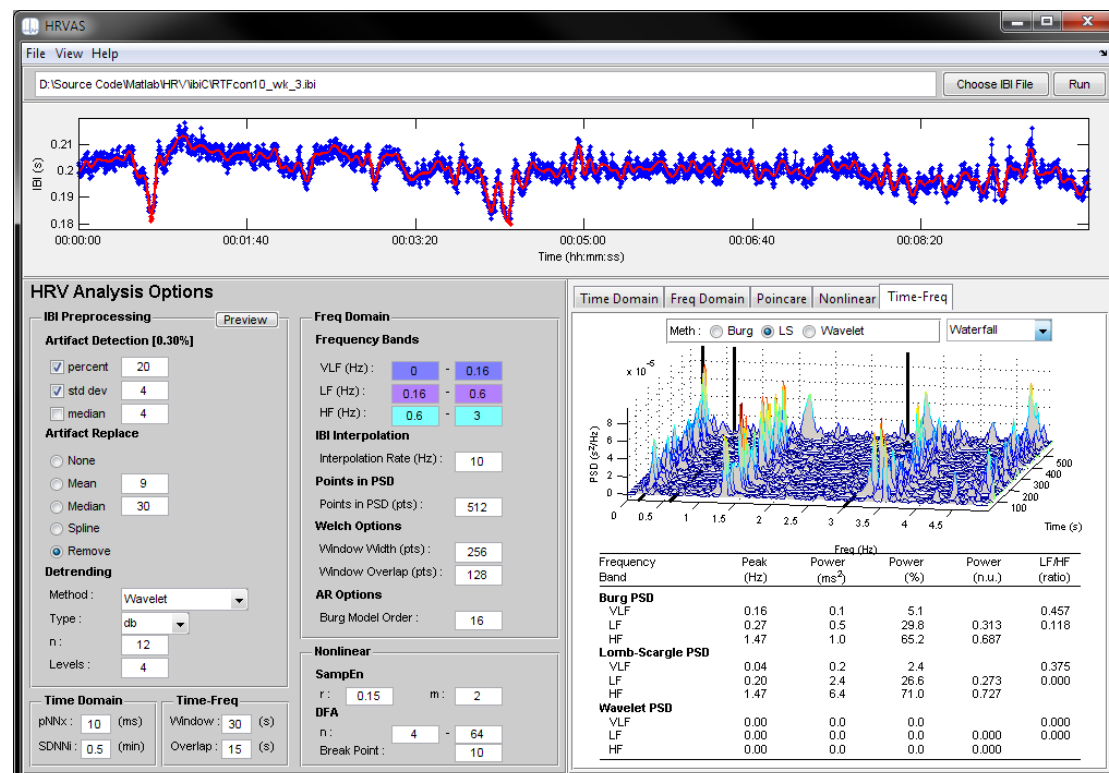


Figure 21: HRVAS visualisation of heart rate activity over time, including breakdowns of instantaneous HR and variability across a number of measures (Ramshur, 2010).

As with the other bio-signals, there are a range of commercially available software that could be used as alternatives, most of which are easier to use and offer many more advanced features to assist in data processing and analysis, such as the ACKHRV component of the AcqKnowledge software by BioPac Systems (BioPac, 2016), which features fully automated processing and analysis of heart rate variability. However, as with the other commercial software previously discussed, the AcqKnowledge software is relatively expensive, rendering it inappropriate for this research project.

The HRVAS MATLAB toolbox is one of the most commonly used HRV processing and analysis tools, offering a good range of features sufficient for the processing and analysis of HR and HRV data in this research project, and has also been successfully used within existing publications (Ramshur, 2010).

### 3.3.1.2.2. IBM SPSS Statistics

SPSS Statistics is a popular software tool that is used for complex statistical and interpretive analysis in social science fields, and allows for the relatively simple examination of highly complicated data sets using features such as descriptive statistics, bivariate statistics, linear regression, and factor analysis. In addition, it also provides features to facilitate data management and documentation (IBM, 2015).

SPSS is a tool that is commonly used in research involving psychophysiology for analysis across multiple data sources and variables (Cacioppo, Tassinary and Berntson, 2007; Mayers, 2013), typically with the data being imported and analysed in-depth after the biosignal data has undergone pre-processing and pre-analysis in specialized software described in previous sections. SPSS is more appropriate than common alternatives such as Microsoft Excel as it offers a greater number of specialized tools for the analysis of complex and multivariate datasets such as those involved in this research project.

### 3.3.2. The Game

Most social video games inextricably combine myriad game components within their core game play experience, which makes it inherently difficult to differentiate between the specific influences that individual components have upon players. Therefore, in order to test and compare the affective and physiological responses to specific components of video games, a video game was required that provides distinct discrete game modes that differentiate between different types of social game play and the components therein. The use of multiple different video games to test different components is not ideal, as this would introduce the game itself as an additional confounding variable within the study. This research project therefore elected to examine a single video game across all aspects of the study.

To adequately address the research objectives of this project, the video game selected needed to meet certain criteria. The selection criteria for video game were as follows:

- Must provide multiple discrete options for both varying social game play, so that components can be measured independently of one another.
- Should have relatively short play times (5-10 minutes) so that each game can be measured through to completion, allowing for the entire game experience to be captured within each experimental condition.
- Must be available on a modern game platform.
- Must be able to replay the same mission/mode multiple times, so that play sessions are consistent between participants.
- Should provide an interesting and fun play experience.
- Should have a Metacritic game rating of at least 70, to help ensure that the selected game is likely popular and of a high quality.

### 3.3.2.1. *Tetris Party Deluxe*

Tetris Party Deluxe met *all* of the selection criteria of this research project:

- It provides multiple distinct social game mode variations, with typical play times of between 5 and 10 minutes.
- The available game modes allow for the individual assessment of varying social game play and individual components therein.
- Tetris is one of the most ubiquitous and popular video games ever made, available on practically all gaming platforms both modern and antiquated.
- Tetris is typically considered to be interesting and fun to play, and has a Metacritic rating of 72 (Metacritic, 2012).

Tetris Party Deluxe on the Nintendo Wii platform (Tetris, 2010) is a modern spin upon the classic Tetris video game, and includes a variety of new mechanics and game modes over the original. The classic Tetris game is one of the most ubiquitous video games ever made, and has been available on virtually every device capable of playing video games since it was first released in 1984. Tetris is an arcade-style puzzle game in which players manipulate geometric shapes, called Tetriminos, of different dimensions by moving them left to right and rotating them in 90 degree increments as they fall down the vertical playing field.

The objective of the game is to create horizontal lines of 10 blocks without gaps, at which point the line will disappear and any blocks above the line will fall down. As the game progresses, the Tetriminos fall faster, making it increasingly difficult for players to correctly place them. The game ends when the stack of Tetriminos reaches the top of the playing field, blocking the fall of new Tetriminos.

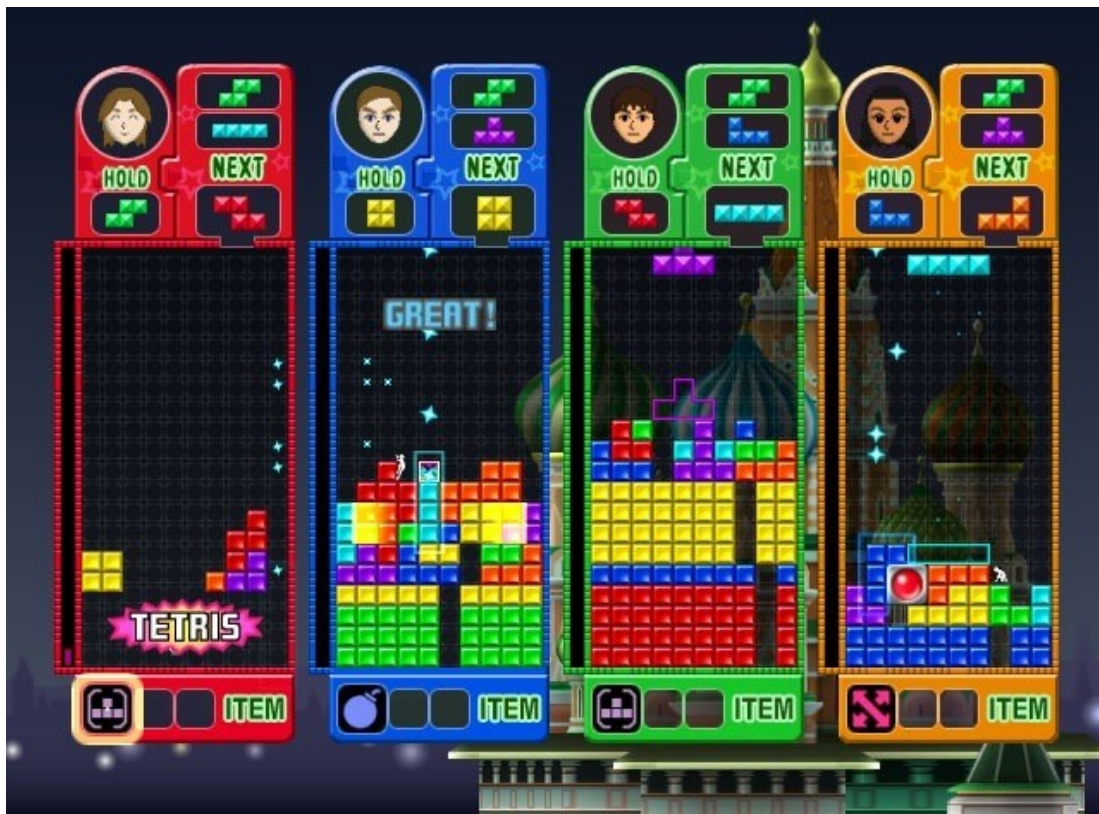


Figure 22: One of the competitive game modes in Tetris Party Deluxe, facilitating direct competition between players (Tetris, 2010).





Figure 23: One of the cooperative game modes in Tetris Party Deluxe, facilitating direct cooperation between players (Tetris, 2010).

Tetris Party Deluxe was released in 2010, and introduces new game mechanics and features, such as competitive and cooperative game modes, as well as additional features such as the ability to use “bombs” or other special items to influence or inhibit the performance of competitors.

Due to its popularity and apparent ubiquity over the past decades, the majority of participants were familiar with Tetris and its basic mechanics, and most required minimal coaching prior to experimentation. Additionally, the shared history of Tetris experience across participants had the potential additional benefit of reducing training effects between participants, as the majority of participants were already proficient at playing the game. Training effects can skew the data when participants initially do not know how to complete a task but then learn and become more proficient at it as the experiment progresses (Suresh, 2011), and so are undesirable within experiments.

As an additional precaution against training effects, participants were all familiarised with the mechanics and control scheme for the game, and were provided with instructions and the rules for each game mode that they were to play immediately prior to playing each mode. These instructions and rules were provided via an instruction sheet rather than via instructions from the researcher, so as to ensure that all participants were given the same information in a consistent manner.

Tetris Party Deluxe proved successful throughout the pilot stages of this research project (see Section 3.6), and so it was decided that it was an appropriate choice for the primary study. A brief discussion of alternative video games considered is provided in Appendix A.

### 3.3.3. Video Recording

To facilitate the effective analysis of game play interactions across social contexts and in terms of in-game events and components, it was necessary to record videos of both the game and of participants as they played the video game. These video recordings allowed for participant and game play interaction observation retroactively, which assisted in the identification of interesting phenomena such as unexpected spikes or dips in psychophysiological signals, as well as providing further information on the physical behaviours of participants, such as gestures, fidgeting, and social interactions.

The game play was recorded using an Avermedia Live Gamer Portable (LGP) device (Avermedia, 2015) attached in-line between the Nintendo Wii and the TV that the game was being displayed on. The LGP was in turn attached to a dedicated recording computer running Windows 8.1 and using the Open Broadcast Software program (OBSProject, 2015), where the game video and audio was recorded alongside other pertinent information.

The participants were recorded using a Microsoft LifeCam HD-3000 device, which was attached to the same Windows 8.1 recording computer as the LGP, with the



audio and video being recorded alongside the game video and audio in the Open Broadcast Software program. This allowed for both sets of audio and video streams to be automatically synchronized at the time of recording, thereby avoiding issues of video streams that do not align during processing and analysis.

### 3.3.4. Baselines

Baseline recordings for all psychophysiological measures were obtained for three minutes both immediately before and immediately after each game play session. During these baselines, participants were asked to quietly watch a video of waves crashing in an ocean, thereby minimizing any potential stimuli that might increase physiological signals above what might otherwise have been their respective norms. These baseline recordings provided what is essentially the current default psychophysiological state of each participant, including their resting heart rate and their current normal level of skin conductivity under the given environmental conditions (Cacioppo, Tassinari and Berntson, 2007; Kivikangas *et al.*, 2011).

The recording of physiological signal baselines is an essential step in research involving psychophysiology, as these are compared against the recordings obtained during experimental conditions to assess the changes produced in response to interactions or events within the video game, if any (Cacioppo, Tassinari and Berntson, 2007; Kivikangas *et al.*, 2011). The measures obtained prior to the play sessions provide information on any residual or lasting impact that the video game may have had upon the psychophysiological state of the participant. By comparing the baselines both before and after the play sessions, data obtained within the play session can be normalised (Cacioppo, Tassinari and Berntson, 2007).

### 3.4. Primary Measures: Physiological

Having established the instruments used within this project, this section now discusses the physiological (primary) measures selected for use, including discussions justifying their selection.

Addressing the research questions of this project requires the use of multiple overlapping physiological measures to assist in the valid analysis and interpretation of results. These include the use of heart rate and galvanic skin response signals for the measurement of arousal responses, and the use of facial electromyography signals for the measurement of valence signals. The following section discusses the selection of these measures.

#### 3.4.1. Brain Activity: Electroencephalography (EEG)

The EPOC+ EEG headset device (discussed in Section 3.3.1.) was capable of recording EEG signals through 16 discrete electrodes placed across the scalp of participants, and is capable of differentiating between five different frequency bands of EEG signals as they occur across the scalp. Previous studies of video game play experiences have correlated beta and gamma signals with physiological arousal and engagement, and alpha frequencies have been associated with relaxation and light mental workloads (Nacke, 2015; Nacke, 2010; Nacke *et al.*, 2010; Ravaja *et al.*, 2005). Due to the nature of video games, these three frequency bands were expected to be the most prevalent within this study.

To ensure that the EEG signals were captured correctly and consistently, electrodes were placed using the 10-20 international standard, which provides a series of 70 locations upon the scalp where electrodes can be placed, and instructions for how to ensure that the electrodes are placed correctly (Trans Cranial Technologies, 2012). Due to the smaller number of electrodes available with the EPOC+ EEG headset device, the electrode placement targeted only those brain regions that have been previously correlated with engagement, attention, emotion, communication, and information processing.

### 3.4.2. Affective Valence: Facial Electromyography (EMG)

The BioRadio device (discussed in Section 3.3.1) can record EMG signals on a series of different locations upon the face of participants, and is capable of providing data on the activation of facial muscles that are used in a wide range of different facial expressions. As previously discussed, EMG signals have previously been found to be a useful and valid instrument in the measurement of the valence of emotions, and therefore EMG constitutes a valuable additional measure against which measures of physiological arousal can be compared.

Measures of arousal (e.g. heart rate and electrodermal activity) can show how active the body and its organs are, which can provide significant insights into how the body is responding to stimuli and events. However, without a measure of valence there is no valid way of knowing whether the participant arousal was positive (e.g. participant is joyous or excited) or negative (e.g. participant is angry or frustrated) (Cacioppo, Tassinary and Berntson, 2007).

The collection of EMG signals also allows for method to assist in the removal of artifacts from the EEG data (often referred to as an EOG), which makes the processing and subsequent analysis of the EEG data quicker, easier, and more reliable (Cacioppo, Tassinary and Berntson, 2007). EMG signals are collected in the same manner as EEG signals, in that involves the placement of electrodes over sites of interest. In this research project, it involved the placement of electrodes over specific facial muscles that are correlated with facial expressions.

To minimise the number of electrodes used, in keeping with the minimalist and naturalistic testing goals of this research project, only two facial muscles were selected for measurement. These two muscles were selected because they were highly correlated with either negative (*Corrugator supercilii*) or positive (*Zygomaticus major*) affective responses in a substantial body of existing literature (Boxtel, 2010). These muscles are briefly discussed below.

### 3.4.2.1. *Corrugator supercilii*

The Corrugator supercilii muscle is responsible for moving the eyebrows downward and inwards, producing a wrinkled and furrowed appearance. It is commonly referred to as the “frowning” muscle, and its activation has been shown across a number of studies to be entirely correlated with negative affective responses such as frustration, sadness, fear, and anger (Boxtel, 2010).



Figure 24: Illustration of facial muscles, with Corrugator supercilii highlighted (Janis *et al.*, 2007).

Due to the decision to limit the number of muscles to be recorded in this project, facial muscles selected needed to be highly correlated with the valences required, so that a minimum number of electrodes could be used to gain the highest value and quality of data. The Corrugator supercilii muscle was determined to be among the most appropriate, as it has been shown to only activate during negatively valenced affective responses to stimuli (Boxtel, 2010; Cowley *et al.*, 2016).

### 3.4.2.2. *Zygomaticus major*

The Zygomaticus major muscle is responsible for moving the angle of the mouth superiorly and posteriorly, producing a smile expression. It is commonly referred to as the “smiling” muscle, and its activation has been shown across multiple studies to be highly correlated with positively valenced affective responses such as happiness, excitement, and joy (Boxtel, 2010).

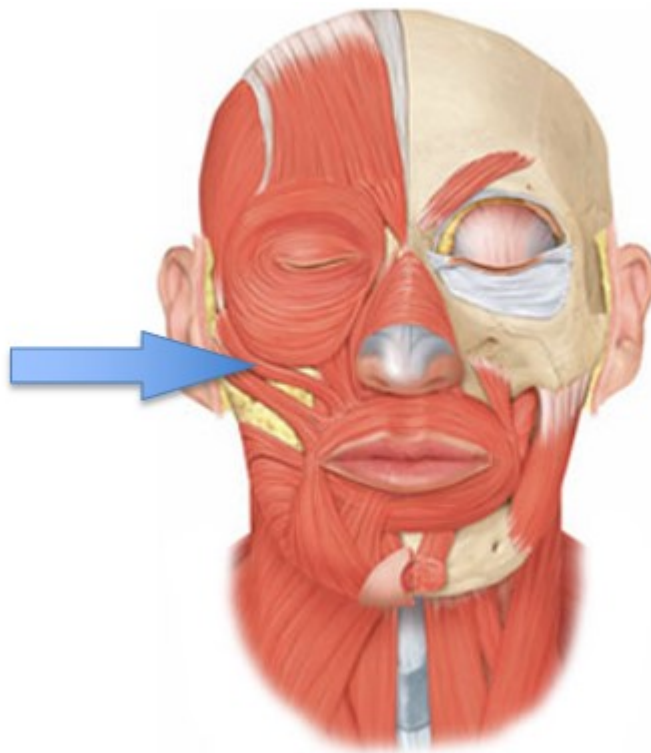


Figure 25: Illustration of facial muscles, with Zygomaticus major highlighted (Janis *et al.*, 2007).

Similarly to the selection criteria for the Corrugator supercilii muscle, the Zygomaticus major muscle was selected for this research project because it has been shown to only activate in the presence of positively valenced affective responses (Boxtel, 2010; Cowley *et al.*, 2016), and so provided an opportunity to record fewer

muscles whilst still reliably capturing the two major poles of negatively vs positively valenced emotions.

### 3.4.3. Physiological Arousal: Heart Rate Photoplethysmography (PPG)

The BioRadio device can record heart rate signals in two ways. The first is through the placement of ECG sensors on the chest to measure electrical activity generated by the heart muscle as it contracts during heart beats, which can provide detailed information on the specific contraction patterns of the heart, and the intervals between different components of the heart beat process (Cacioppo, Tassinary and Berntson, 2007).

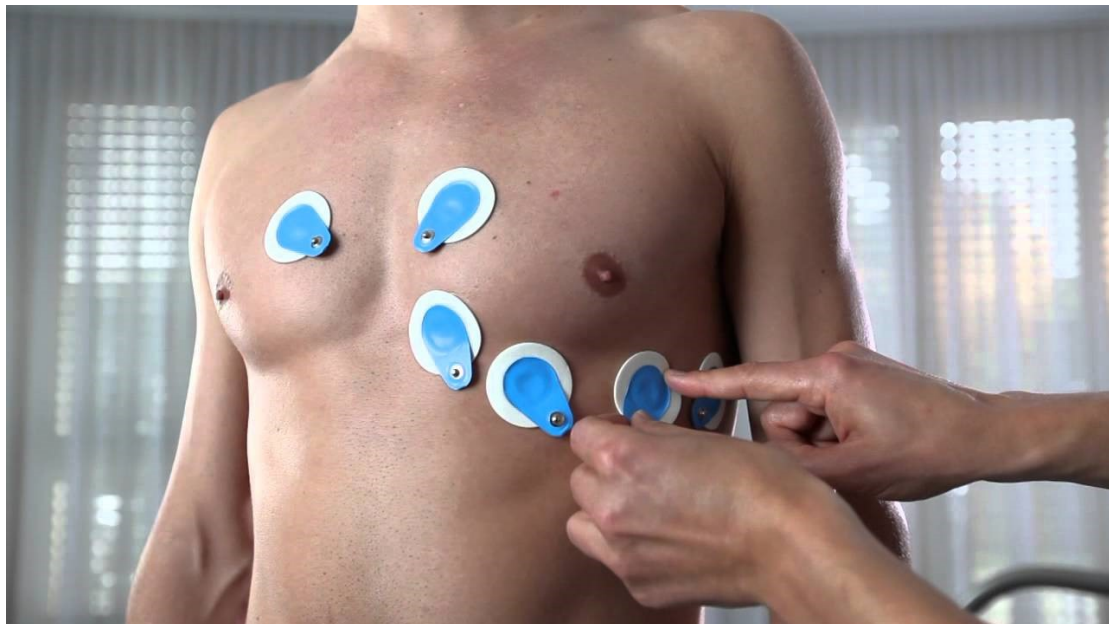


Figure 26: Example of ECG electrode placement on participant.

The second way is through using a photoplethysmography (PPG) sensor on a finger to measure the heart beat and oxygen saturation of the blood stream, and it works by directing specific wavelengths of light into the finger and then measuring the wavelengths of the light that then exits the finger. This allows the device to determine the heart rate and the oxygen saturation of the blood through measuring the pulse of arterial blood in the finger (Cacioppo, Tassinary and Berntson, 2007).



The heart rate and level of blood oxygenation have been strongly correlated with physiological arousal across multiple domains (Cacioppo, Tassinary and Berntson, 2007), with heart rate being the most popular measure for physiological arousal due both to the availability and low cost of devices that are capable of measuring it. As with other psychophysiological measures of arousal, the instantaneous heart rate is unable to determine the valence of physiological arousal, and can simply show variations in physiological arousal (Cacioppo, Tassinary and Berntson, 2007). The analysis of heart rate data therefore requires additional measures, such as EMG, to determine the valence of responses.



Figure 27: Pulse Oximeter photoplethysmography (PPG) placement on participant.

While ECG can provide more comprehensive information on the specific patterns and rhythm of the heart muscle, it is more susceptible to interference and artefacts that may arise due to extraneous muscle movements in the arms and torso (Cacioppo, Tassinary and Berntson, 2007). In clinical or laboratory environments, this does not represent much of an issue, however within a naturalistic social game environment it is likely that the ECG signal will become highly unusable due to natural body movements and gestures that form part of social interactions.

Fortunately, the PO method of measuring the heart rate is less susceptible to extraneous body movements, due to the way in which it measures heart signals. The PO sensor is also extremely easy to apply, as it simply clips onto the end of any finger. Therefore, it was decided that while the PO sensor is not a perfect technology, it was the most appropriate device for the recording of cardiac activity as a measure of physiological arousal within this research project.

### **3.4.4. Physiological Arousal: Galvanic Skin Response (GSR)**

The BioRadio device can record GSR signals through the use of two electrodes placed close together either on the hand, fingers, or forearm of participants. Also commonly referred to as Electrodermal Response (EDR) or Electrodermal Activity (EDA), GSR measures the amount of sweat present on the surface of the skin by determining the conductivity of the skin surface (Cacioppo, Tassinary and Berntson, 2007). The activation of sweat glands due to physiological arousal results in the skin surface containing more moisture, which in turn increases the electrical conductivity of the skin surface (Cacioppo, Tassinary and Berntson, 2007). The GSR signal is therefore a simple yet effective measure that is commonly used as a measure of physiological arousal, and is commonly used in conjunction with heart rate signals to strengthen and corroborate the determination of physiological arousal.



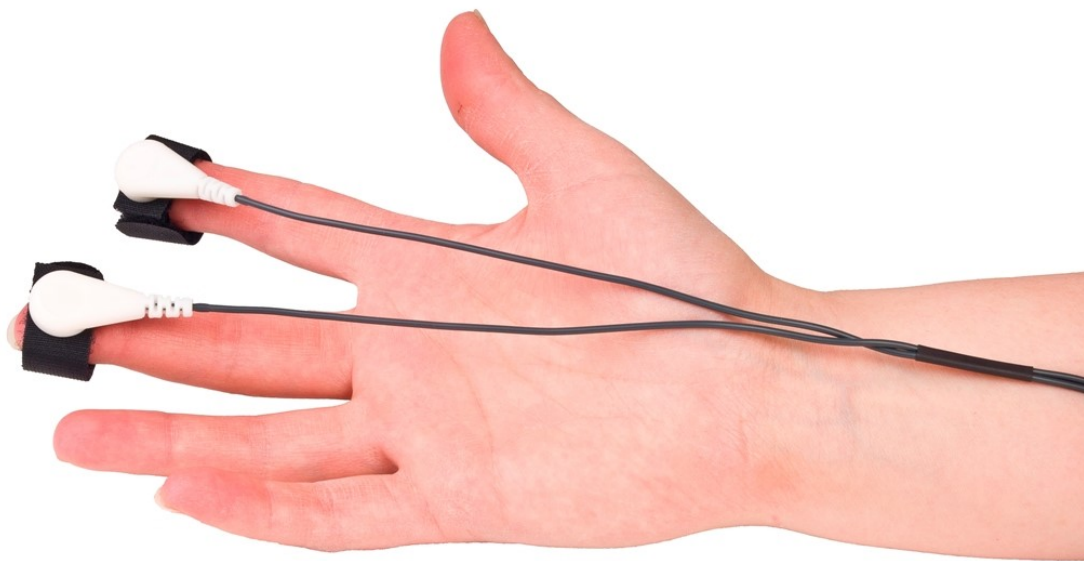


Figure 28: GSR electrode placement on participant.

The flexible placement of the GSR sensors means that it is not likely to disturb players during experimental sessions, and the nature of the GSR measurement means that it is not susceptible to extraneous body movement.

#### **3.4.5. Primary Measures: Summary**

This section has provided an overview of the primary, physiological, measures selected for use within this research project. These have been selected for their ability to contribute towards the objectives and questions of this project, primarily their correlation with affective responses. The following section discusses the selection of secondary measures that enrich the objective physiological measures with additional contextual data, and facilitate improved insights for the purposes of addressing the research questions.

### 3.5. Secondary Measures: Observational and Subjective Data

The use of subjective surveys has been shown to be important within objective studies of psychophysiology, as they provide researchers with evidence from multiple perspectives with which to triangulate factors and corroborate results (Nacke, Grimshaw, and Lindley 2010).

To capture a sufficient range of contextually rich data, many additional data points, including subjective survey instruments and observations, were selected for use within this research project. A major consideration in the selection and rejection of instruments was the time required to undertake the experimental process, as it was deemed undesirable for participants to have to spend excessive time filling in surveys during and after experimental conditions due to potential fatigue. With this goal, only those surveys considered essential to the aims of the research were included.

#### 3.5.1. Subjective Measures of Player Experiences: Game Experience Questionnaire (GEQ)

The Game Experience Questionnaire (GEQ) consists of 7 different components, each of which includes between 5 to 6 questions regarding different aspects of the experience of playing a video game (e.g. “I lose track of where I am”), which are assessed through a 5-point Likert scale that ranges between 1: “Not At All” to 5: “Extremely” (Norman, 2013; Ijsselstein, de Kort & Poels, 2008).

The 7 components of the GEQ aim to assess multiple components of the player experience within video games, including:

1. **Flow:** A rewarding mental state where the player is fully engaged, focused, and feeling high levels of enjoyment (Ijsselstein, de Kort & Poels, 2008).
2. **Tension:** The experience of mental or emotional strain, which can include the feeling of conflict, pressure, or anxiety (Ijsselstein, de Kort & Poels, 2008).

3. **Challenge:** The experience of how difficult or challenging activities were within the game, which could range between extremely easy and extremely hard (Ijsselstein, de Kort & Poels, 2008).
4. **Negative Affect:** The experience of negatively valenced emotions, such as frustration, fear, or anger (Ijsselstein, de Kort & Poels, 2008).
5. **Positive Affect:** The experience of positively valenced emotions, such as joy, excitement, or happiness (Ijsselstein, de Kort & Poels, 2008).
6. **Presence/Immersion:** The experience of immersion within the game, which considers how well the game held players' attention, as well as how involved players were emotionally (Ijsselstein, de Kort & Poels, 2008; Rigby and Ryan, 2007).
7. **Competence:** The experience of proficiency, which includes how capable or skilled they feel within the game (Ijsselstein, de Kort & Poels, 2008; Rigby and Ryan, 2007).

The GEQ survey is one of the most commonly used subjective instruments in research involving the measurement of the experiences players have while playing video games, with a wealth of extant literature illustrating its validity and reliability (e.g. Cowley *et al.*, 2014; Norman, 2013; Kokil, 2013; Nacke and Lindley, 2009). Due to the prevalence of the GEQ in existing video game research, along with the survey covering the key aspects of the player experience that this project is interested in, such as negative and positive affect, it was decided that the GEQ was the most appropriate survey for the measurement of subjective video game play experience.

A brief discussion of alternative subjective instruments considered but rejected for the measurement of player experiences are presented in Appendix B.

### 3.5.2. Subjective Measures of Affective Responses:

#### 3.5.2.1. *Video recorded facial expressions*

As an additional, contextually rich, data source of the affective responses that participants experienced during video game play, video recordings were made of facial expressions, behaviours, and interactions of participants during experimental sessions. This provided an avenue for the retrospective analysis of facial expressions, which could be used to provide an additional data point to triangulate and compare with the physiological signals recorded, facilitating more rigorous analysis of visible affective responses and social interactions.

Video recordings of facial expressions are discrete and non-invasive, with participants experiencing little to no discomfort because of the recordings. This makes video recording a valuable additional measure, although it does have significant drawbacks in terms of the time required to process the video recordings (if completed manually), which typically are conducted frame by frame and may require between 5 and 50 hours of analysis per hour of video recorded. Nonetheless, this enabled for more advanced analysis of the second hypothesis, as it provides a way of assessing facial expressions produced under similar game conditions both with and without the presence of the recording electrodes.

This provided valuable insights into possible moderating influences that the presence of psychophysiological equipment had upon video game experiences, with the potential to provide evidence to either provide evidence for or against the use of psychophysiology as a discrete objective measure of affective responses within future video game studies.

#### 3.5.2.2. *Self-Assessment Manikin (SAM):*

In addition to the use of recorded facial expressions, the secondary measurement of post-stimulus affective responses can be completed using the Self-Assessment Manikin (SAM) self-report survey instrument. The SAM was first developed in 1985

by Lang (1985) with the intent of measuring emotional responses to advertising materials (Morris and Karrh, 1994), and since then has been used and validated within several multimedia fields, including video game research (Ravaja *et al.*, 2004; Ivory and Kalyanaraman, 2007; Klarkowski *et al.*, 2016). SAM relies upon visual representations of emotions upon vaguely humanoid characters, and requires individuals to select where they believe that their current emotional state fits within three dimensions of pleasure, arousal, and dominance (PAD), a model of affective responses first proposed by Russell and Mehrabian (1977).

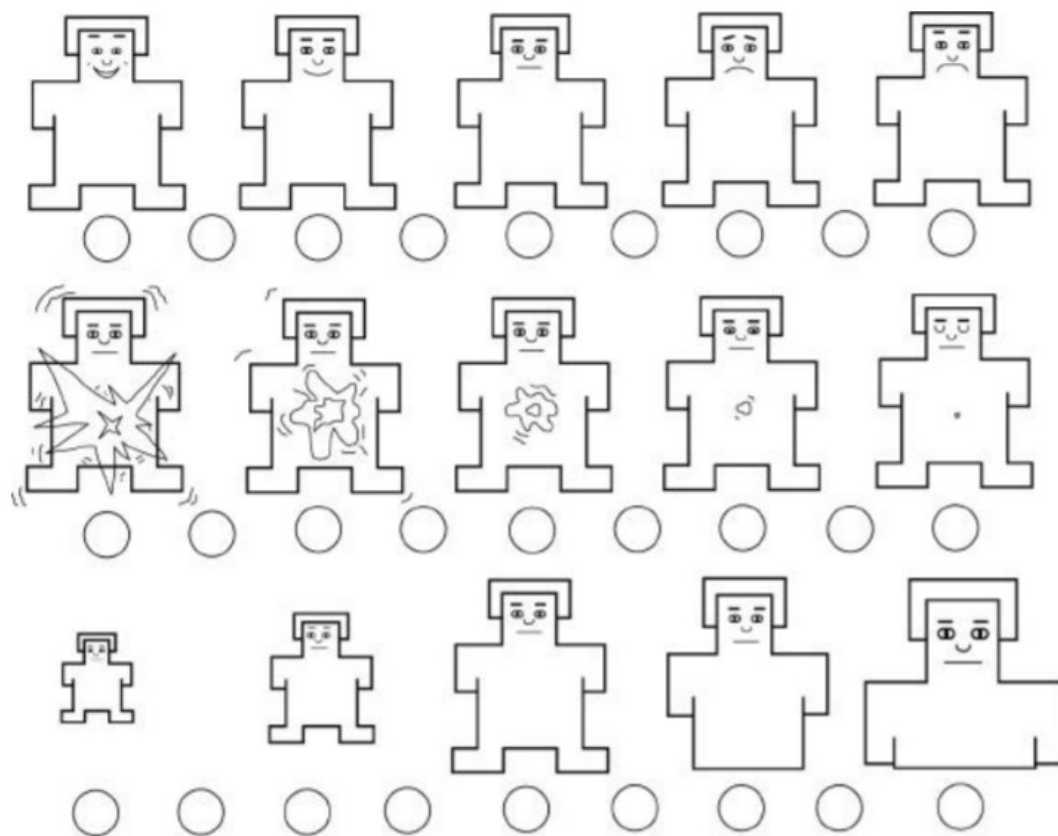


Figure 29: The Self-Assessment Manikin (SAM) from the IAPS norms (Lang, Greenwald, and Bradley, 1989). The rows (from top to bottom) require individuals to rate their perceived pleasure, arousal, and dominance.

The use of simple stylised representations has been found to improve the consistency of participant results, and unlike other available instruments has even proven reliable across cultures as it does not rely upon interpretation of questions

(Morris, 1995). As an additional benefit, due to the participant only having to provide three responses per test, the SAM is very fast to administer.

The SAM is therefore considered an ideal instrument for the secondary measuring of affective responses to video game stimuli, being easy to administer, simple to interpret, and has been used and found valid and reliable within previous studies within the video game domain (Ivory and Kalyanaraman, 2007; Nacke, Nacke and Lindley, 2009; Klarkowski *et al.*, 2016). A brief discussion of an alternative subjective instrument considered is presented in Appendix C.

### 3.5.3. Demographic Survey

A demographic survey was used to collect basic information about participants, including their age, gender, their average time spent playing video games per week, preferences for video game genres, hobbies and interests, as well as their area of occupation or study. This information was collected through free-form text fields, allowing participants to enter their information without restrictions or and avoiding biases introduced through leading questions. This demographic information enhanced the ability for this research project to identify potential individual factors that can be correlated with specific affective and physiological responses whilst playing video games, and may exert moderating influences on these responses.

### 3.5.4. Secondary Measures Summary

This section has discussed the selection of secondary, additive, measures that enrich the primary physiological data and better facilitate the addressing of the objectives and research questions of this project. Having selected both the primary and secondary measures necessary for addressing these objectives and questions, the following sections discuss the design and conduct of three pilot studies that were conducted over the course of this project.

### 3.6. Methodology: Pilot Studies

Prior to committing to a final study design, procedures, and instruments, a series of pilot studies were first conducted with the intention of testing key methodological assumptions and attempting to identify problems early on. This section details the design and conduct of these pilots, including an overview of their results.

#### 3.6.1. Method and Research Design

This research project attempts a novel methodological approach, to investigate affective responses to events occurring during video game play. While there have been psychophysiological investigations into social and multimedia phenomena within other fields, the novelty of this project and investigation results in a notable lack of methodological antecedents within the domain of video game research. Broadly, this study intends to reflect experimental procedures and recommendations made by Mandryk *et al.* (2006; 2007), Nacke (2007; 2009; 2011), Drachen *et al.* (2010), Ravaja *et al.* (2005; 2008), and Kivikangas *et al.*, (2011).

However, standard practice within the field of HCI research is to ensure that the study design and associated procedures are appropriate and able to reliably provide data directly relating to the research questions through the conductance of a series of pilot tests, usually intended to test various aspects of the experimental design and procedure. The results of these pilots have been used to inform the design and procedure for the primary study, the aims of which have been discussed previously.

#### 3.6.2. Participant Recruitment

Participants for the pilot studies were recruited via word of mouth from within the University of Tasmania, and respondents were asked to recruit from their friends and family to gain some broader demographics for the pilot experiments. Participants were required to be over the age of 18, and have no significant understanding of the experimental aims or procedures prior to participation.

Each of the two pilot studies included either 8 or 10 participants, with anyone participating in earlier pilots being eligible to also participate in subsequent pilots. Individuals who participated in any of the pilots were not eligible to participate in the primary study itself, as a measure to help counter against any possible biases. The total number of participants across all four pilots was 36 (14 female, mean age 26.86 years (SD  $\pm$  4.79)).

### **3.6.3. Pilot 1: Basic physiological arousal through Heartrate**

A small scale ( $n = 8$ ) pilot experiment was conducted, with the aim of identifying whether a consumer-grade heart rate monitor is capable of measuring variations in participant heart rate elicited through interaction with a video game. It was intended that this would verify the possibility of objectively measuring differences in physiological arousal across different video game conditions, though no specific hypothesis was put forward as to the result.

An additional objective was to check whether the selected video game, Tetris Party Deluxe, was sufficient to elicit physiological responses, and by extension to verify its suitability for use within this project. Pilot 1 was also intended to provide preliminary insights into how participants respond to the instruments and preliminary experimental process, which were then used to guide subsequent pilots as well as the resulting primary study of this project.

#### **3.6.3.1. Pilot 1 Procedure**

Pilot 1 included 3 experimental conditions in a Within-Participant design (Solo Play, Competitive Play, Cooperative Play) which were measured in a randomised order for all participants. A baseline heart rate was recorded prior to the commencement of each condition.

Participants were provided with an introductory overview of the research project, including being informed of what game they were to play, which was Tetris Party



Deluxe (Tetris, 2010), on the Nintendo Wii platform. Participants were informed that they would be playing non-socially, competitively, and cooperatively. Participants were informed of the nature of the testing device, which consisted of a Polar H7 Bluetooth Heart Rate Sensor connected to an iPhone device running the DigiFit fitness application for data recording purposes. Participants were provided with information sheets describing the experiment, and asked to sign informed consent forms if they were willing to proceed with participation.

Participants were shown how to correctly place and wear the heart rate strap, which sits immediately below the breasts, and then asked to put on the strap in private. A baseline measure was then recorded for 2 minutes, the purpose of which was to obtain the average heart rate at rest for each participant, as this varies between individuals. After the baseline was recorded, participants were provided with instructions on how to play the game, and were shown the preparatory screen immediately prior to the game starting.

Participants were assigned to conditions randomly, and were asked to proceed through them one at a time. After the completion of each condition, which last for approximately 8 minutes each, players were allowed 3 minutes to allow their heart rate to normalise, at which point an additional baseline was recorded for 2 minutes. This procedure was repeated until all three conditions had been completed.

### **3.6.3.2. Pilot 1 Data Processing**

Keeping with established protocols for the assessment of physiological data (Cacioppo, Tassinary, and Berntson, 2007; Nacke *et al.*, 2011), the heart rate data for each participant was visually inspected to ensure that there were no artefacts or corrupted data segments.

The nature of physiological data means that it has varying baseline values between individuals, making it difficult to compare values between across participants. The standard practice of analysis of heart rate data is therefore to transform it into a

percentage of variance from baseline, whereby the value of the heart rate above baseline value for each participant is taken as a percentage (Nacke, 2015; Cacioppo, Tassinary, and Berntson, 2007). For example, where before a participant data might have had a heart rate of 86 beats per minute, the normalised data would now read as a 6% variance. This renders the data comparable across participants, and simplifies the analysis and interpretation processes.

The data was then exported to SPSS for statistical analysis, including both descriptive and inferential statistics, which are discussed in the proceeding section.

### 3.6.3.3. Pilot 1 Results

The within-subject study design used in Pilot 1, with a single independent variable with three levels of depth (Competitive Play, Cooperative Play, Solo Play), a single dependent variable (Heart Rate). As the study design had a single continuous dependent variable, and the independent variable contained three levels, a one-way within-subjects ANOVA was selected as an appropriate method of analysis (Girden, 1992; Kirk, 2013).

Descriptive Statistics			
	Mean	Std. Deviation	N
HR_Solo	9.22	2.79	8
HR_Coop	13.56	2.21	8
HR_Comp	12.78	2.90	8

*Table 2: Descriptive statistics for Pilot 1, showing mean data +- standard deviation.*

One-way within-subjects ANOVAs have three additional assumptions that must first be tested prior to proceeding. Firstly, there should be no significant outliers within the dataset. Outliers can have a negative impact on the one-way ANOVA by distorting the differences between levels of independent variables, and can reduce generalisability of the sample results (Ghosh and Vogt, 2012).

As evidenced through inspection of a boxplot, there were no outliers in the dataset (see Figure 30 below).

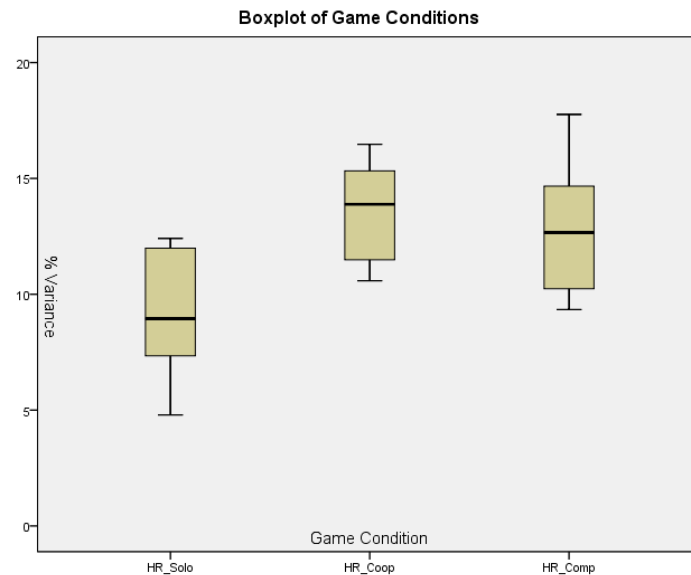


Figure 30: Boxplot used to assess for outliers, with none evidenced.

Secondly, the dataset for the dependent variable should be normally distributed. Despite being considered robust to violations of normality, ANOVAs require approximately normally distributed data to produce valid results (Girden, 1992).

Heart rate data was normally distributed within each game condition, as assessed by Shapiro-Wilk's test ( $p > .05$ ) (see Table 3 below).

Tests of Normality						
	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
HR_Solo	.192	8	.200*	.928	8	<b>.499</b>
HR_Coop	.186	8	.200*	.931	8	<b>.525</b>
HR_Comp	.131	8	.200*	.944	8	<b>.655</b>

\*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Table 3: Shapiro-Wilk's Test of Normality, evidencing normal distribution.

Lastly, there should be equal variances of differences between all levels of the independent variable, referred to as sphericity. Violating the assumption of sphericity increases incidences of Type 1 error, where significant results are reported when there isn't one (Weinfurt, 2000).

Mauchly's test of sphericity indicated that the assumption of sphericity has not been violated,  $\chi^2(2) = 2.183$ ,  $p = .336$ .

### Mauchly's Test of Sphericity

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.
Arousal	.695	2.183	2	.336

Table 4: Mauchly's Test of Sphericity, evidencing that sphericity has not been violated.

With all three assumptions met, a one-way within-subjects ANOVA was conducted. As illustrated in Table 5 below, Play conditions elicited statistically significant changes in Heart Rate,  $F(2, 14) = 16.315$ ,  $p < .0005$ , *partial*  $\eta^2 = .7$ .

### Tests of Within-Subjects Effects

Measure: HR

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Arousal	85.511	2	42.755	16.315	.00022	.700
Error(Arousal)	36.689	14	2.621			

Table 5: Output of one-way within-subjects ANOVA, evidencing a significant difference in Heart Rate between Play conditions.

To further investigate the significant effects found in the one-way within-subjects ANOVA, a pairwise comparison between Play conditions within Pilot 1 was conducted using a post hoc test. A post hoc test compares all possible combinations of the levels of the independent variable, in this case multiple paired-samples t-tests with a Bonferroni adjustment (Maxwell and Delaney, 2004).

### Pairwise Comparisons with Bonferroni adjustment

Measure: HR

(I) Arousal	(J) Arousal	Mean Difference (I-J)	Std. Error	Sig. <sup>b</sup>	95% Confidence Interval for Difference <sup>b</sup>	
					Lower Bound	Upper Bound
HR_Solo	HR_Coop	-4.335*	.912	.0046	-7.186	-1.484
	HR_Comp	-3.560*	.917	.0181	-6.428	-.692
HR_Coop	HR_Solo	4.335*	.912	.0046	1.484	7.186
	HR_Comp	.775	.542	.5867	-.919	2.469
HR_Comp	HR_Solo	3.560*	.917	.0181	.692	6.428
	HR_Coop	-.775	.542	.5867	-2.469	.919

Based on estimated marginal means

\*. The mean difference is significant at the .05 level.

Table 6: Multiple paired-samples t-tests with a Bonferroni adjustment, evidencing statistically significant differences in means between HR\_Solo and HR\_Cooperation Play conditions.

Data are mean +- standard deviation, unless otherwise stated. There was an increase in Heart Rate from 9.22% +- 2.79% in the Solo Play condition to 13.56% +- 2.21% in the Cooperative Play condition, a statistically significant increase of 4.335% (95% CI, 1.483 to 7.186),  $p < 0.05$ .

There was also an increase in Heart Rate from 9.22% +- 2.79% in the Solo Play condition to 12.78% +- 2.9% in the Competitive Play condition, however it was not a statistically significant increase at 3.56% (95% CI, .692 to 6.428),  $p = .13$ .

The difference in Heart Rate between the Competitive Play and Cooperative Play conditions was smaller, with a statistically insignificant increase from 12.78% +- 2.9% to 13.56% +- 2.21%, an increase of .775% (95% CI, -.919 to 2.469),  $p = .587$ .

#### 3.6.3.4. Pilot 1 Summary

The results of Pilot 1 demonstrated that even with a very small sample size of 8 participants, statistically significant differences with moderate effect sizes can be

found in the heart rate and therefore general level of arousal between different social game play conditions. This confirmed both that there are objectively measurable differences in arousal through simply varying social contexts such as competition and cooperation, and that Tetris Party Deluxe was sufficiently able to elicit such responses.

This data corroborated the results of previous studies within the video game domain, where it was found that heart rate can be made to vary across varying social or non-social conditions (Nacke *et al.*, 2009; Drachen, L. Nacke, *et al.*, 2010; Clark, 2012). As heart rate is a key predictor of affective arousal (Cacioppo, Tassinari and Berntson, 2007; Cowley *et al.*, 2016), the ability to measure variances in it across game conditions was confirmed as a data source of interest for the remaining pilot studies, as well as the primary study within this project.

### 3.6.4. Pilot 2: EEG brain activity

A small scale ( $n = 10$ ) pilot experiment was conducted, with the aim of identifying whether electroencephalography (EEG) is a viable measure within social video game environments. EEG is a valuable measure capable of providing insights into brain activity during multimedia experiences, including assessing cognitive pathways and processes, and providing metrics for affect at high-temporal resolution (Wehbe & Nacke, 2013; Mandryk & Nacke, 2016; Cacioppo and Tassinari, 1990; Cacioppo, Tassinari and Berntson, 2007; Salminen and Ravaja, 2007; Astolfi *et al.*, 2010; Baillet, 2011).

EEG is known to be highly susceptible to errors due to body movements, including talking, and so studies using EEG typically require participants to sit still and not speak, to reduce artefacts and data contamination (Cacioppo, Tassinari and Berntson, 2007; Burgess, 2013). However, this requirement would compromise the environmental context of social video game play and potentially reduce the applicability of the resulting data to normal video game contexts.

The objectives of this project include generating insights into affective responses within naturalistic video game contexts, and so Pilot 2 was undertaken to assess the reliability of EEG measures within a natural social video game environment. To this end, the EPOC+ EEG headset was selected as it was a lightweight wireless device with relatively fast setup, and is relatively non-intrusive compared to alternative EEG instruments.

### **3.6.4.1. Pilot 2 Procedure**

Pilot 2 included 3 experimental conditions using a Within-Participant design (Solo Play, Competitive Play, Cooperative Play), which were measured in a Latin Square randomised order for all participants. A two-minute baseline EEG was recorded prior to the commencement of each condition.

Participants were first provided with an introduction to the research project, including being informed of the game they were to play, which was Tetris Party Deluxe (Tetris, 2010), on the Nintendo Wii platform. Participants were informed that they would be playing non-socially, competitively, and cooperatively. Participants were informed of the nature of the testing device, which consisted of an Emotiv EPOC+ EEG Headset (Emotiv, 2015), connected to a computer running Windows 8.1 and the Xavier software package for data recording purposes. Participants were provided with information sheets describing the nature of the experiment and instruments used, and asked to sign informed consent forms if they were willing to proceed with participation.

Participants then had the EPOC+ device prepared and placed upon their scalps, and the device was calibrated using the Xavier software. Participants were then asked to sit quietly for two minutes while a baseline measure was recorded, the purpose of which was to obtain the average brain activity at rest for each participant, as this varies between individuals. After the baseline was recorded, participants were provided with instructions on how to play the game, and were shown the preparatory screen immediately prior to the game starting.

Participants were assigned to conditions randomly, and were instructed to proceed through them one at a time. After the completion of each condition, which took approximately 8 minutes each, players were allowed 3 minutes for their brain activity to normalise, at which point an additional baseline was recorded for 2 minutes. This procedure was repeated until all three conditions had been completed.

### 3.6.4.2. Pilot 2 Data Processing

Pilot 2 used an Emotiv EPOC+ (Emotiv, 2015) headset to assess the viability of using electroencephalography (EEG) as a continuous real-time measure of brain activity during social video game play. There was a total of three experimental conditions, namely Solo Play, Cooperative Play, and Competitive Play. Additionally, a two-minute baseline EEG recording was made prior to each experimental condition.

The EEG data was cleaned and processed using two freely available software packages. The first, EEGLAB (Delorme and Makeig, 2004; Brunner, Delorme and Makeig, 2013), is a popular open source MATLAB toolbox that is used for processing and analysing EEG data. The second, Brainstorm (Mosher and Baillet, 2005; Tadel *et al.*, 2011), is an open source standalone program that is dedicated to the analysis of MEG and EEG data, and offers some additional tools for the inference of brain activity and functional connectivity between brain regions.

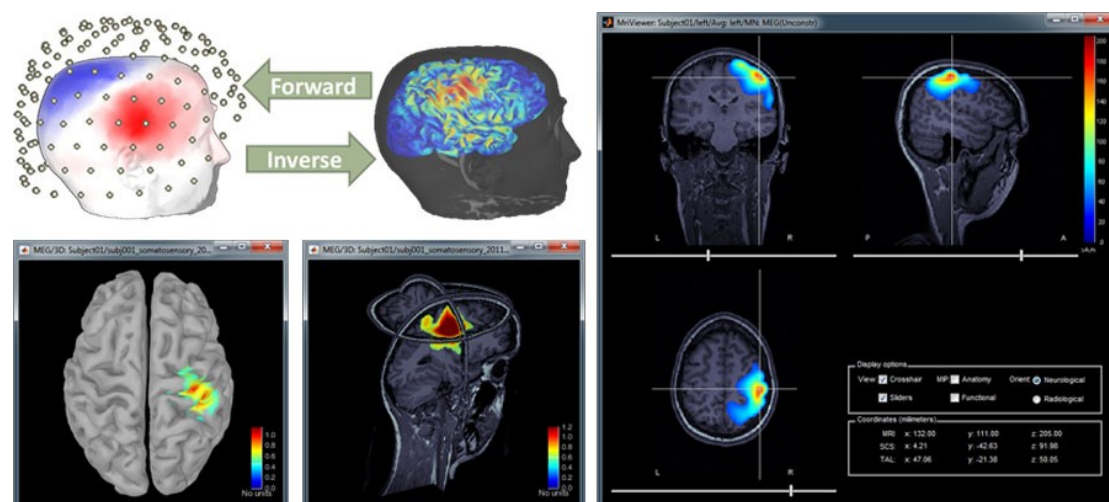




Figure 31: Visualisation of Event Related EEG data in Brainstorm, illustrating region of brain activation and functional connectivity.

A spectral analysis of the EEG data was undertaken so that the five EEG frequency bands could be determined from the single EEG trace. A fast-Fourier Transform (FFT) and spectral analyses were undertaken to determine the strength of each frequency band any given length of time (Watling, 2012).

The brain activity frequency bands and their associated mental states (Nacke *et al.*, 2009; Lim and Reeves, 2010):

- Delta Band (0.1 - 3 Hz): Associated with deep sleep or meditative trances.
- Theta Band (4 - 7 Hz): Associated with being drowsy, inactive, or idle mental workloads.
- Alpha Band (8 - 12 Hz): Associated with being calm or relaxed, or moderate mental workloads.
- Beta Band (12 - 30 Hz): Associated with being alert, engaged, and focused.
- Gamma Band (35 - 45 Hz): Associated with actively processing information and intense mental workloads.

Of the five EEG frequency bands, only four were of potential interest to this study, as the delta frequency band is usually only found during deep sleep or meditation, and so this band was discarded.

The EEG trace for each experimental condition was measured in ten-second epochs. This is shorter than the epochs that are typically used within EEG studies, however because of the social nature of the second pilot study, a shorter epoch length was considered appropriate to limit data loss through artefacts (Cacioppo, Tassinary and Berntson, 2007; Kivikangas *et al.*, 2011; Liu, Sourina and Nguyen, 2011; Cowley *et al.*, 2016).

There are several extraneous influences that can compromise EEG recordings, such as talking, swallowing, and blinking (Cacioppo, Tassinary and Berntson, 2007; Cowley

*et al.*, 2016). The resultant artefacts increase the relative powers across the EEG frequency bands, and as such it was necessary to visually inspect each epoch prior to conducting data analysis. Where an epoch contained 4 or more seconds of artefacts, the epoch was deleted and recorded as missing data. This process was repeated for each epoch recorded within the pilot.

Due to the social nature of the second pilot, with participants being allowed to speak and move as they might usually during game play, approximately 70% of the EEG data was discarded in this way prior to analysis. The mean relative powers of every epoch for each frequency band was then exported to SPSS for statistical analysis, including both descriptive and inferential statistics, which are discussed in the proceeding section.

#### 3.6.4.3. Pilot 2 Results

A within-subject study design used in Pilot 1, with a single independent variable with three levels of depth (Competitive Play, Cooperative Play, Solo Play), a single dependent variable (brain activity per frequency band). A one-way within-subjects ANOVA was selected as an appropriate method of analysis (Laerd Statistics, 2015).

Descriptive Statistics			
	N	Mean	Std. Deviation
Solo 4-7Hz	71	.2183	.12231
Solo 8-12Hz	71	.2718	.13143
Solo 12-30Hz	71	.1251	.07249
Solo 35-45Hz	71	.1590	.09507
Comp 4-7Hz	71	.1260	.07001
Comp 8-12Hz	71	.2001	.12194
Comp 12-30Hz	71	.2238	.12770
Comp 35-45Hz	71	.2492	.15102
Coop 4-7Hz	71	.0983	.06350
Coop 8-12Hz	71	.2697	.14555
Coop 12-30Hz	71	.1823	.11433
Coop 35-45Hz	71	.2215	.12121

Table 7: Descriptive statistics for Pilot 2, showing mean data  $\pm$  standard deviation.

As with Pilot 1, the data was first inspected to ensure that the three assumptions required for one-way ANOVAs were met. Firstly, as evidenced through inspection of a boxplot, there were no outliers in the dataset (see Figure 32 below).

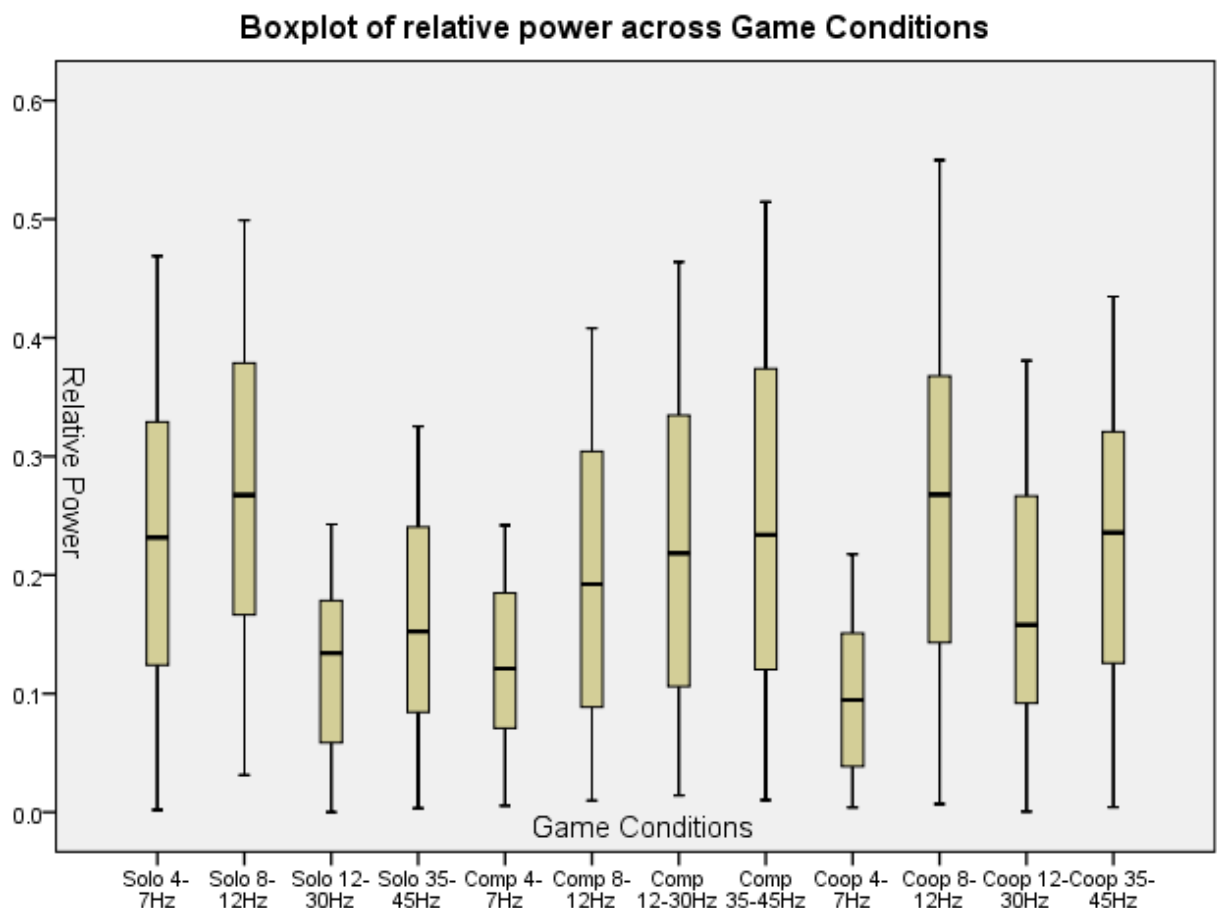


Figure 32: Boxplot used to assess for outliers, with none evidenced.

Secondly, as evidenced through Shapiro-Wilk's Test of Normality, the relative EEG power dataset was normally distributed within each game condition,  $p > 0.05$ . Lastly, as evidenced through Mauchly's Test of Sphericity, the principle of sphericity has been violated,  $\chi^2(2) = 14.255$ ,  $p < 0.01$ . A Greenhouse-Geisser adjustment was made (Greenhouse and Geisser, 1959) to the measures, correcting the violation.

With two assumptions met and the third corrected for, a one-way within-subjects ANOVA was conducted for each of the four EEG frequency bands.

#### 3.6.4.3.1. Relative Theta Activity

As illustrated in Table 8 below, Play conditions elicited statistically significant changes in relative Theta band brain activity,  $F(1.685, 117.979) = .321, p < 0.01$ ,  $partial \eta^2 = .321$ .

**Theta Band: Tests of Within-Subjects Effects**

Measure: Relative\_Theta

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Theta	.541	1.685	.321	33.098	<b>.0071</b>	.321
Error(Theta)	1.145	117.979	.010			

Table 8: Output of one-way within-subjects ANOVA for relative Theta frequency activity, evidencing a significant difference across the Play conditions.

To further investigate the significant differences in relative Theta band activity found in the one-way within-subjects ANOVA, a pairwise comparison between Play conditions within Pilot 2 was conducted using a post hoc test, in this case multiple paired-samples t-tests with a Bonferroni adjustment (Maxwell and Delaney, 2004).

**Theta Band: Pairwise Comparisons**

Measure: Relative\_Theta

(I) Theta	(J) Theta	Mean Difference (I-J)	Std. Error	Sig. <sup>b</sup>	95% Confidence Interval for Difference <sup>b</sup>	
					Lower Bound	Upper Bound
Solo_Theta	Comp_Theta	<b>.092*</b>	.016	<b>.006</b>	.053	.132
	Coop_Theta	<b>.117*</b>	.017	<b>.005</b>	.075	.160
Comp_Theta	Solo_Theta	<b>-.092*</b>	.016	<b>.006</b>	-.132	-.053
	Coop_Theta	<b>.025</b>	.012	<b>.105</b>	-.004	.053
Coop_Theta	Solo_Theta	<b>-.117*</b>	.017	<b>.005</b>	-.160	-.075
	Comp_Theta	<b>-.025</b>	.012	<b>.105</b>	-.053	.004

Based on estimated marginal means

\*. The mean difference is significant at the .05 level.

Table 9: Multiple paired-samples t-tests with a Bonferroni adjustment, evidencing statistically significant differences in means.

Data are mean  $\pm$  standard deviation, unless otherwise stated. There was a decrease in relative Theta from  $.218 \pm .122$  in the Solo Play condition to  $.126 \pm .07$  in the Competitive Play condition, a statistically significant decrease of  $.092$  (95% CI,  $.053$  to  $.132$ ),  $p < .05$ .

There was also a decrease in relative Theta from  $.218 \pm .122$  in the Solo Play condition to  $.098 \pm .063$  in the Cooperative Play condition, a statistically significant decrease of  $.117$  (95% CI,  $.075$  to  $.16$ ),  $p < .05$ .

There was a small difference in relative Theta from  $.126 \pm .07$  in the Competitive Play condition to  $.098 \pm .063$  in the Cooperative Play condition, but it is statistically insignificant difference of  $.025$  (95% CI,  $-.004$  to  $.053$ ),  $p = .105$ .

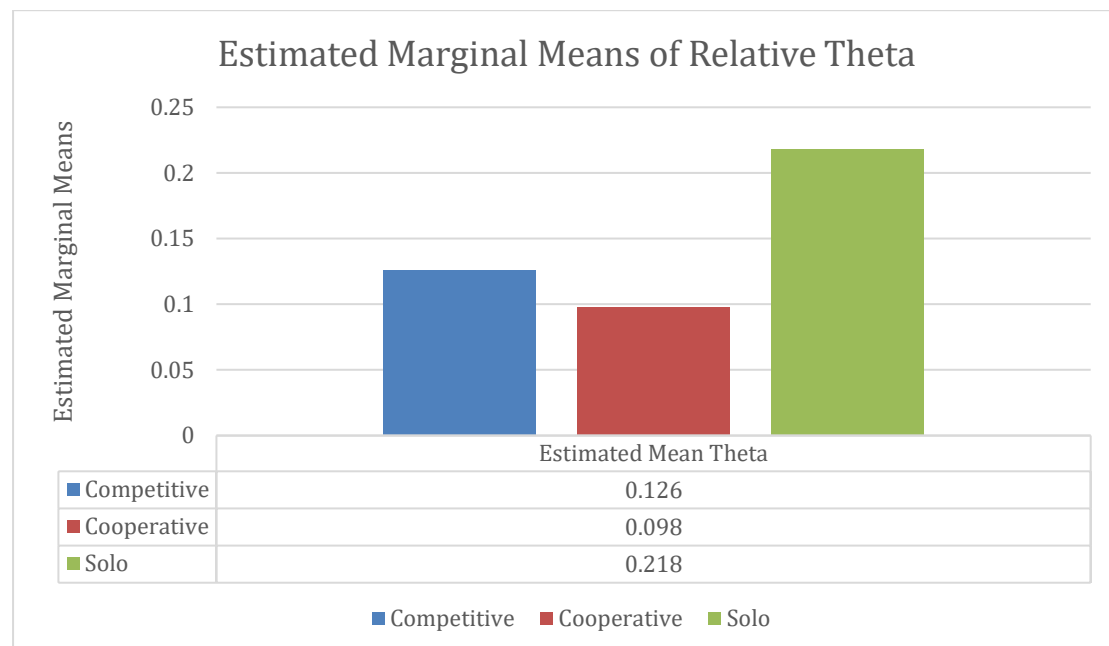


Figure 33: Mean differences in Theta activity between Competitive Play, Cooperative Play, and Solo Play.

### 3.6.4.3.2. Relative Alpha Activity

As illustrated in Table 10 below, Play conditions elicited statistically significant changes in relative Alpha band brain activity,  $F(1.93, 135.11) = .123$ ,  $p < 0.005$ ,  $\text{partial } \eta^2 = .089$ .

#### Theta Band: Tests of Within-Subjects Effects

Measure: Relative\_Alpha

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Alpha	Greenhouse-Geisser	.237	1.930	.123	6.860	.002	.089
Error(Alpha)	Greenhouse-Geisser	2.415	135.109	.018			

Table 10: Output of one-way within-subjects ANOVA for relative Alpha frequency activity, evidencing a significant difference across the play conditions.

To further investigate the significant differences in relative Alpha band activity found in the one-way within-subjects ANOVA, a pairwise comparison between Play conditions in Pilot 2 was conducting using a post hoc test, in this case multiple paired-samples t-tests with a Bonferroni adjustment (Maxwell and Delaney, 2004).

#### Alpha Band: Pairwise Comparisons

Measure: Relative\_Alpha

		Mean Difference	Std. Error	Sig. <sup>b</sup>	95% Confidence Interval for Difference <sup>b</sup>	
(I) Alpha	(J) Alpha	(I-J)			Lower Bound	Upper Bound
Solo_Alpha	Comp_Alpha	.072*	.020	.002	.021	.122
	Coop_Alpha	.002	.024	.664	-.057	.061
Comp_Alpha	Solo_Alpha	-.072*	.020	.002	-.122	-.021
	Coop_Alpha	-.070*	.022	.006	-.122	-.017
Coop_Alpha	Solo_Alpha	-.002	.024	.664	-.061	.057
	Comp_Alpha	.070*	.022	.006	.017	.122

Based on estimated marginal means

\*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Table 11: Multiple paired-samples t-tests with a Bonferroni adjustment, evidencing statistically significant differences in means.

Data are mean  $\pm$  standard deviation, unless otherwise stated. There was a decrease in relative Alpha from  $.272 \pm .131$  in the Solo Play condition to  $.200 \pm .122$  in the Competitive Play condition, a statistically significant difference of  $.072$  (95% CI,  $.021$  to  $.122$ ),  $p < .005$ .

There was also a decrease in relative Alpha from  $.269 \pm .146$  in the Cooperative Play condition to  $.200 \pm .122$  in the Competitive Play condition, a statistically significant difference of  $.07$  (95% CI,  $.017$  to  $.122$ ),  $p < .005$ .

The difference in relative Alpha band activity between Solo Play and Competitive Play conditions was very small, with a statistically insignificant difference of  $.002$  (95% CI,  $.057$  to  $.061$ ),  $p = .665$ .

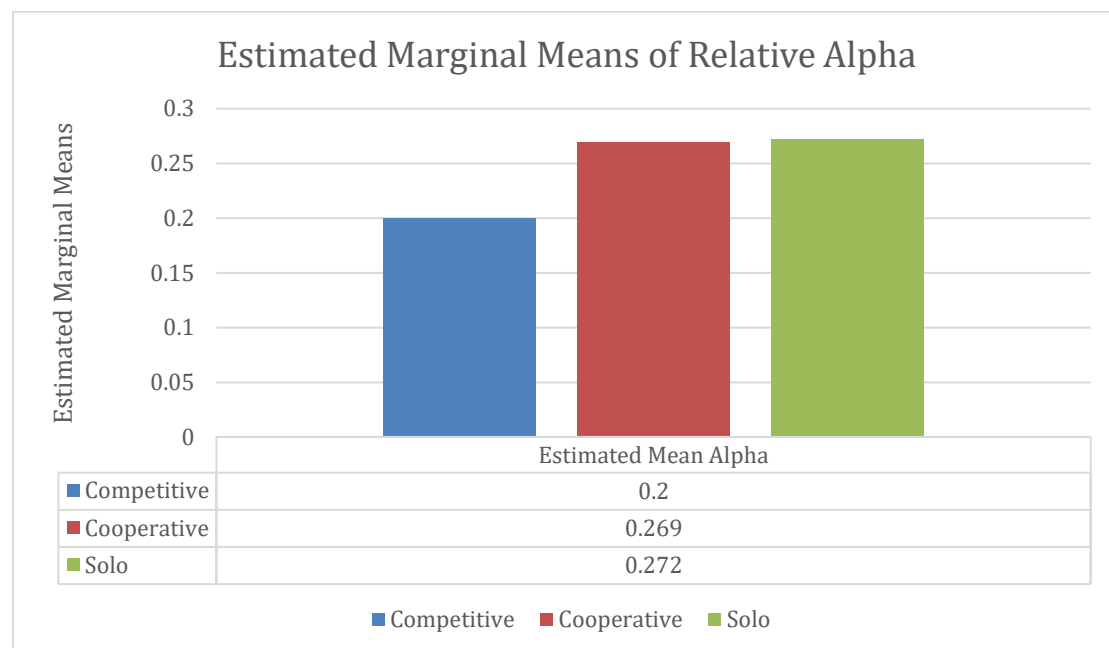


Figure 34: Mean differences in Alpha activity between Competitive Play, Cooperative Play, and Solo Play.

### 3.6.4.3.3. Relative Beta Activity

As illustrated in Table 12 below, Play conditions elicited statistically significant changes in relative Beta band brain activity,  $F(1.756, 122.89) = .18, p < 0.005$ , *partial*  $\eta^2 = .18$ .

**Beta Band: Tests of Within-Subjects Effects**

Measure: Relative\_Beta

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Beta	Greenhouse-Geisser	.349	1.756	.199	15.372	.003	.180
Error(Beta)	Greenhouse-Geisser	1.589	122.893	.013			

Table 12: Output of one-way within-subjects ANOVA for relative Beta frequency activity, evidencing a significant difference across Play conditions.

To further investigate the significant differences in relative Beta band activity found in the one-way within-subjects ANOVA, a pairwise comparison between Play conditions within Pilot 2 was conducted using a post hoc test, in this case multiple paired-samples t-tests with a Bonferroni adjustment (Maxwell and Delaney, 2004).

**Beta Band: Pairwise Comparisons**

Measure: Relative\_Beta

(I) Beta	(J) Beta	Mean Difference (I-J)	Std. Error	Sig. <sup>b</sup>	95% Confidence Interval for Difference <sup>b</sup>	
					Lower Bound	Upper Bound
Solo_Beta	Comp_Beta	-.099*	.017	.002	-.141	-.057
	Coop_Beta	-.057*	.015	.001	-.095	-.020
Comp_Beta	Solo_Beta	.099*	.017	.002	.057	.141
	Coop_Beta	.042	.021	.150	-.010	.093
Coop_Beta	Solo_Beta	.057*	.015	.001	.020	.095
	Comp_Beta	-.042	.021	.150	-.093	.010

Based on estimated marginal means

\*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

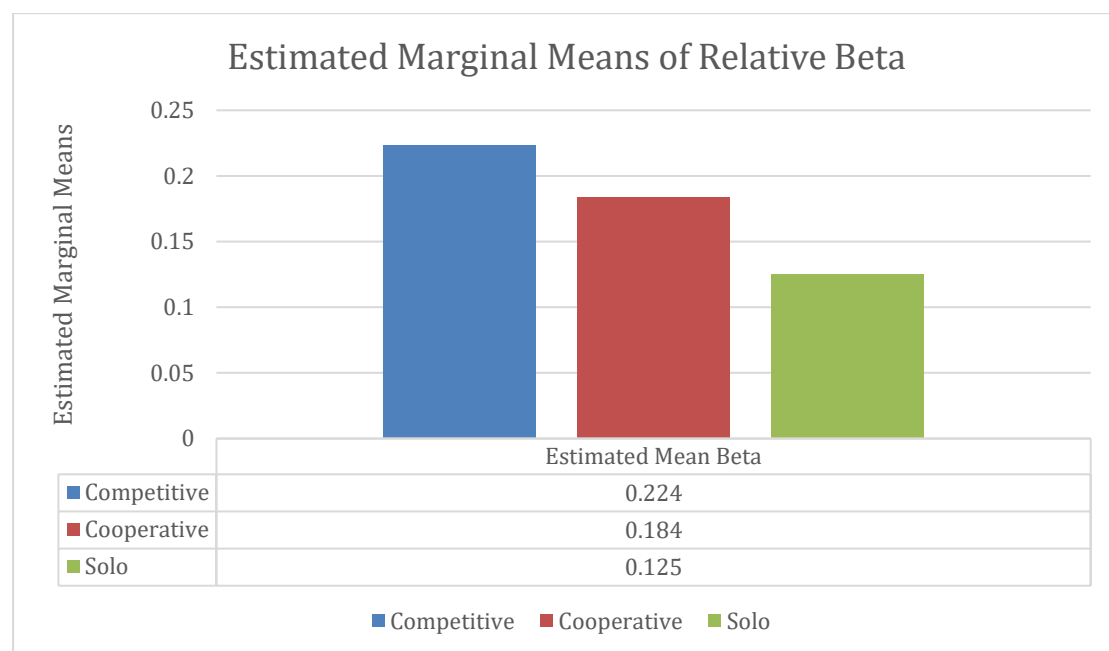
Table 13: Multiple paired-samples t-tests with a Bonferroni adjustment, evidencing statistically significant differences in means.



Data are mean  $\pm$  standard deviation, unless otherwise stated. There was an increase in relative Beta from  $.125 \pm .072$  in the Solo Play condition to  $.224 \pm .128$  in the Competitive Play condition, a statistically significant increase of  $.099$  (95% CI,  $.057$  to  $.141$ ),  $p < 0.005$ .

There was also an increase in relative Beta from  $.125 \pm .072$  in the Solo Play condition to  $.182 \pm .114$  in the Cooperative Play condition, a statistically significant difference of  $.057$  (95% CI,  $.02$  to  $.095$ ),  $p < 0.005$ .

Although there is a small decrease in relative Beta activity between Competitive Play and Cooperative Play conditions, these are statistically insignificant at  $.042$  (95% CI,  $-.01$  to  $.093$ ),  $p = .15$ .



*Figure 35: Mean differences in Beta activity between Competitive Play, Cooperative Play, and Solo Play.*

## 3.6.4.3.4. Relative Gamma Activity

As illustrated in Table 14 below, Play conditions elicited statistically significant changes in relative Gamma band brain activity,  $F(1.95, 136.4) = .156, p < 0.01$ .

**Gamma Band: Tests of Within-Subjects Effects**

Measure: Relative\_Gamma

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Gamma	Greenhouse-Geisser	.303	1.949	.156	10.519	<b>.007</b>	.131
Error(Gamma)	Greenhouse-Geisser	2.017	136.398	.015			

Table 14: Output of one-way within-subjects ANOVA for relative Gamma frequency activity, evidencing a significant difference across Play conditions.

To further investigate the significant differences in relative Gamma band activity found in the one-way within-subjects ANOVA, a pairwise comparison between Play conditions within Pilot 2 was conducted using a post hoc test, in this case multiple paired-samples t-tests with a Bonferroni adjustment (Maxwell and Delaney, 2004).

**Gamma Band: Pairwise Comparisons**

Measure: Relative\_Gamma

(I) Gamma	(J) Gamma	Mean Difference (I-J)	Std. Error	Sig. <sup>b</sup>	95% Confidence Interval for Difference <sup>b</sup>	
					Lower Bound	Upper Bound
Solo_Gamma	Comp_Gamma	<b>-.090*</b>	.021	<b>.002</b>	-.143	-.037
	Coop_Gamma	<b>-.062*</b>	.019	<b>.004</b>	-.108	-.017
Comp_Gamma	Solo_Gamma	<b>.090*</b>	.021	<b>.000</b>	.037	.143
	Coop_Gamma	<b>.028</b>	.020	<b>.521</b>	-.022	.077
Coop_Gamma	Solo_Gamma	<b>.062*</b>	.019	<b>.004</b>	.017	.108
	Comp_Gamma	<b>-.028</b>	.020	<b>.521</b>	-.077	.022

Based on estimated marginal means

\*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Table 15: Multiple paired-samples t-tests with a Bonferroni adjustment, evidencing statistically significant differences in means.

Data are mean  $\pm$  standard deviation, unless otherwise stated. There was an increase in relative Gamma from  $.159 \pm .095$  in the Solo Play condition to  $.25 \pm .151$  in the Competitive Play condition, a statistically significant increase of  $.09$  (95% CI,  $.037$  to  $.143$ ),  $p < .002$ .

There was an increase in relative Gamme from  $.159 \pm .095$  in the Solo Play condition to  $.221 \pm .121$  in the Cooperative Play condition, a statistically significant increase of  $.062$  (95% CI,  $.017$  to  $.108$ ),  $p = .017$ .

There was a small difference between the Competitive and Cooperative Play conditions, however this was a statistically insignificant difference of  $.028$  (95% CI,  $-.077$  to  $.022$ ),  $p = .521$ .

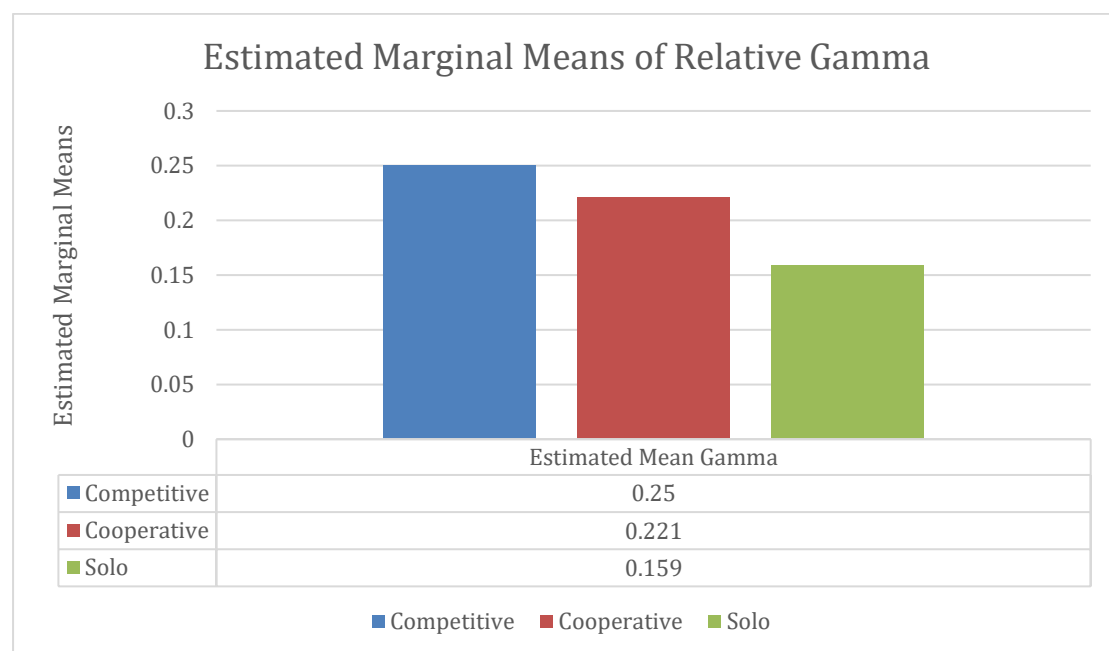


Figure 36: Mean differences in Gamma activity between Competitive Play, Cooperative Play, and Solo Play.

### 3.6.4.4. Pilot 2 Summary

The results of Pilot 2 demonstrated that even with a small sample size of 10 participants, statistically significant differences can be found in EEG brain activity recordings and therefore a general understanding of brain activity frequencies can

be attained through playing varying social game play conditions. This supports the argument that there are objectively measurable cognitive differences between social game play conditions, such as competition and cooperation, and that Tetris Party Deluxe was sufficiently able to elicit such responses.

Participants playing video games within the Solo condition, where they were playing against an AI and did not have any social interactions available to them, were found to have significantly higher levels of relative Theta band brain activity. Theta band is commonly associated with inactivity, drowsiness, and low mental workloads (Cacioppo, Tassinary and Berntson, 2007; Nacke, Stellmach and Lindley, 2011; Cowley *et al.*, 2016), suggesting that playing Tetris in non-social contexts is less stimulating than playing with friends in either Competitive or Cooperative contexts.

To lend additional weight to this finding, the levels of both Beta band and Gamma band brain activity was significantly lower in the Solo condition than in either the Competitive and Cooperative conditions. Beta brain activity is commonly associated with concentration and engagement, while Gamma brain activity is commonly associated with intense mental workloads (Cacioppo, Tassinary and Berntson, 2007; Nacke *et al.*, 2009; Cowley *et al.*, 2016). One possible explanation is that the AI opponent in the Solo Play condition did not require the same degree of concentration or attention, potentially due to being too easy. Klarkowski *et al.* (2016) reported that physiological arousal is directly influenced by challenge afforded by the game, and Nacke *et al.* (2011) reported that cognitive responses appeared to be related to difficulty, particularly where the challenge offered was either too easy or too hard for the player.

Further investigation using recorded gameplay videos revealed that players in the Solo Play condition did not perform any better on average than they did against real players in the Competitive Play condition. This makes it unlikely that the differences in brain activity are related to challenge or difficulty, although it is possible that participants perceived the difficulty to be different, possibly due to

additional stress or expectations associated with playing against friends (Mandryk and Inkpen, 2004; Ravaja *et al.*, 2005; Clark, 2012; Jarvela *et al.*, 2014).

The brain activity measured within Competitive and Cooperative Play conditions are largely similar across all but one frequency bands, where Competitive Play resulted in significantly lower Alpha band activity than both Solo and Cooperative play conditions. Alpha band frequencies are commonly associated with being relaxed, calm, and having low mental workloads (Cacioppo, Tassinary and Berntson, 2007; Cowley *et al.*, 2016). This may suggest that participants experienced higher levels of cognitive arousal or stress when playing the Competitive Play condition compared to both the Solo and Cooperative Play conditions, where they were either competing against an AI in the case of Solo, or cooperating with a friend to complete a challenge in the case of Cooperative.

Despite presenting an interesting avenue for exploratory research, the objective of Pilot 2 was to assess the effectiveness of EEG as a tool within the larger research project. Further investigation into this observation was deemed beyond the scope of Pilot 2, and has been relegated to future research (see Section 6.3).

While the results clearly illustrate that interesting and potentially meaningful insights could be derived through use of EEG within social video game contexts, there were a myriad of issues associated with its use. Most importantly, due to the social interactions and other bodily movements of players, over 70% of all EEG data had to be discarded due to artefacts and other data corruption associated with movements. Prior to discarding the corrupted data, a range of artefact correction techniques were applied in an attempt to minimise data loss, however these ultimately proved insufficient.

The objective of this research project is to provide real-time measures and insights, through which events occurring within video games can be understood in terms of their respective influences upon affect. It was therefore initially intended that Pilot 2

would provide real-time measures on game events, however due to the overwhelming volume of data lost through artefacts, there were too many gaps in the data to provide reliable insights into events in real-time. The EEG data was therefore processed and analysed as a condition average, which while useful and interesting for use with certain research objectives, does not contribute towards the objectives of this project.

Event-Related Potentials (ERP) EEG could have the potential to circumvent issues of social interactions and body movement through the presentation of key stimuli and known points in time, and recording responses to these stimuli (Cacioppo, Tassinari and Berntson, 2007). This approach has proven useful in a range of other multimedia fields, including investigations into the influence of violent media upon cognitive responses (Bartholow, Bushman and Sestir, 2006; Kirmizi-Alsan *et al.*, 2006; Cacioppo, Tassinari and Berntson, 2007; Salter *et al.*, 2012). However, ERP studies require removing the stimuli from its context. In the case of a video game, testing a stimulus from outside of the game environment is not likely to provide the same results as testing it within the game environment, as video games are contextually rich, and breaking them into pieces would fundamentally alter the experience.

It was therefore decided that while EEG might be a useful measure for video game studies, it is not suitable for this project due to the data loss caused by social interactions, which are required for this project to achieve its objectives. EEG was therefore removed as a measure from subsequent studies within this project.

### 3.6.5. Pilot 3: Trial of primary procedure

After the finalization of the psychophysiological instruments and experimental variables to be used within the primary study, a small scale (n=10) experiment was conducted with the aim of testing the full experimental design and procedure, including all prospective subjective instruments. This pilot was intended to identify

any potential issues or complications that could occur within the primary study so that these could be rectified prior to commencing the primary study.

After this third pilot, it was concluded that there were no evident issues with the experimental procedures, measures, or data collected. It was therefore decided that this pilot should proceed as the primary study, with the participant pool and data collected merged into the primary. As all data had been merged and included with the primary experiment, no discussion of them will occur here. Please refer to Section 3.7. for full details of the Primary Experiment, and Chapter 4 for results.

### **3.6.6. Methodology: Pilot Study Summary**

Prior to conducting the primary experiment of this research project, a series of pilot studies were first undertaken to develop and test equipment and experimental conditions.

The pilot study sections of this thesis have reported on the design of these antecedent studies, and provided a brief overview of the results including a discussion of their respective implications as they relate to the overall research project.

### **3.7. Methodology: Main Study**

This research project utilises a repeated-measures (within-participant) experimental design, using a psychophysiological instrument (BioRadio device, by GLNeuroTechnologies) to measure multiple body signals (galvanic skin response (GSR)), facial electromyography (EMG), heart rate photoplethysmography (PPG)) that are highly correlated with affective and physiological responses to multimedia stimuli. Additionally, several survey instruments (Game Experience Questionnaire (GEQ), Self-Assessment Manikin (SAM), and a demographic survey) were used within this study, to provide additional data with which to triangulate and strengthen the results obtained from the objective-quantitative psychophysiological data, and to

assist in the establishment of ground-truth of the physiological results and subsequent inferences.

The primary study was conducted using a 1v2 condition design, with a total of two experimental conditions (Equipment On vs Equipment Off) plus a control baseline, where the presence of equipment was manipulated within participants. This design and set of experimental conditions have been selected as they are capable of drawing out insights into how and why varying events occurring during video game play result in affective responses, and may elucidate other potentially relevant factors influencing the play experience. The 1v2 design facilitates the examination of game aspects both independently for individual effects, as well as interdependently within different variations and combinations. Psychophysiological processes are used to explain these effects, including the specific conditions wherein they occur.

To minimise the influence that the researcher had upon the measurement of psychophysiological signals, the researcher left the room and monitored it and the equipment remotely during game play measurements, and so limited potential confounding social interactions between the participants and the researcher.

For further details on the procedure, analysis, and results of the primary study, refer to Sections 3.7, 3.8, 3.9, 4.0, and 5.0 respectively.

### **3.7.1. Method and Research Design**

The experimental design used for the primary study within this research project was the result of continual evaluation, redevelopment, and refinement based both on the results and lessons learned from the three pilot experiments, and with reference to academic literature from across the domains of video game and psychophysiology (see Chapter 2).

The primary study utilised a repeated-measures (within-participant) 1v2 variable design, with a two-level independent variable of Equipment On vs. Equipment Off.



This split-level condition allowed for events occurring during game play to be identified and analysed post-experiment, but also facilitated the examination of measurement effects associated with the manipulation of psychophysiological equipment.

To control for potential training effects or other extraneous variables (Suresh, 2011), one participant is randomly assigned to a starting condition (either Equipment On or Equipment Off), and the other participant is assigned to the opposite condition to that based on the rules discussed previously. When aggregating the data for all participants, any extraneous training effects should therefore be evenly distributed across all conditions, and thereby reduce skews or biases within the data.

### 3.7.2. Participant Recruitment

For the main experiment within this project a total of fifty-six (56) participants were recruited from members of the Tasmanian public aged 18 or over. Participants were recruited through advertising via email, including bulk emails distributed throughout the University of Tasmania and other organisations, through the placement of adverts on bulletin boards, through radio interviews and adverts, as well as through social media outlets such as Facebook. Respondent volunteers were required to bring a friend or family member with them, so that all participants undertook the experiment as a pair with someone that they were familiar with. There were no other criteria restricting selection.

Roughly equal numbers of both males and females were recruited (30 males, 26 females). Participants had a mean age of 28.7 (+/- 5.4) years, and on averaged reported playing 14.2 (+/- 7.6) hours of video games per week. The recruitment of a broad demographic was targeted gather a participant sample that could be considered at least somewhat representative of the average Australian video game player (47% female, average age of 33, 10 hours per week playing games) (Brand, Lorentz and Mathew, 2014).

Based on previous research using psychophysiology (Mandryk and Inkpen, 2004; Mandryk, Atkins and Inkpen, 2006; Ravaja and Kivikangas, 2008b; Nacke *et al.*, 2009; Drachen, L. Nacke, *et al.*, 2010; Jarvela *et al.*, 2014), it was determined that the primary study required at least 30 participants to have a reasonable chance of gaining statistical significance, and would greatly benefit from including a greater number of participants. The project was successful in recruiting a total of 56 participants, however the data for 12 participants had to be removed from the study due to data corruption during experimental sessions. This resulted in a total of 44 participants remaining in the study.

### 3.7.3. Experimental Setup and Procedures

The following sections describe the setup of the experimental spaces, including overviews of the hardware used, and the experimental procedures used within the primary experiment.

#### 3.7.3.1. Experimental Room Setup

The primary study took place on the Sandy Bay campus at the University of Tasmania. The experimental room was outfitted with a 40-inch flat screen TV, Nintendo Wii gaming console, computers for recording video and bio-signals, and comfortable couch for participants. Participants sat next to each other on the couch, facing the TV.

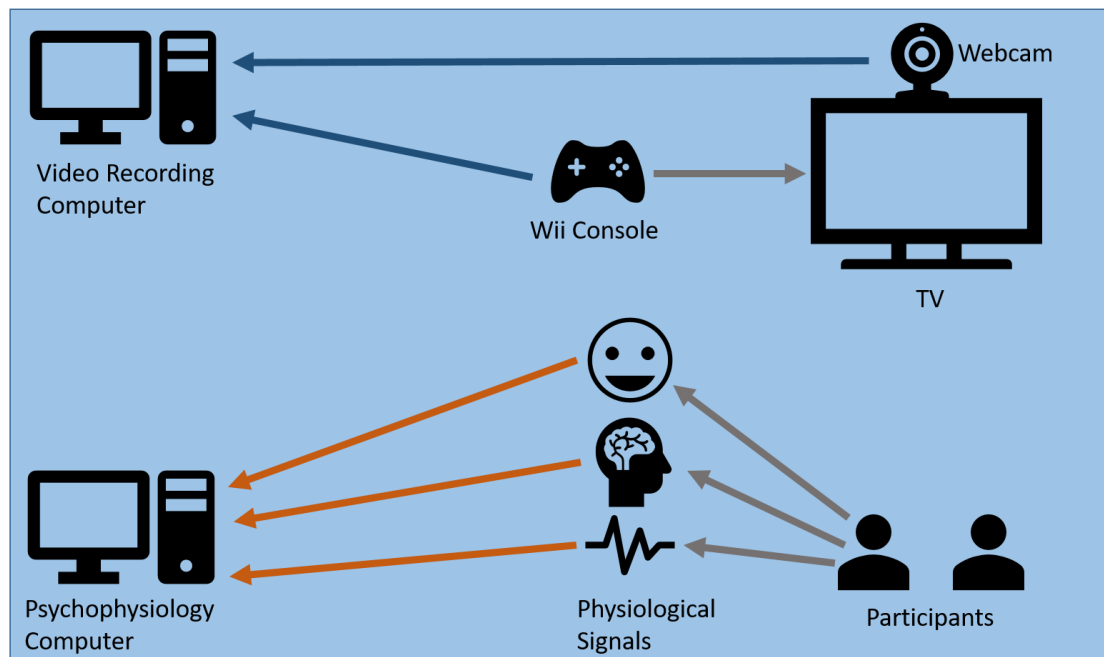


Figure 37: Diagram of experimental room setup.

### 3.7.3.2. Computing Infrastructure

The experimental space was setup with two desktop computers running Windows 8.1 Professional for the purposes of recording data during experimental sessions. During experimental sessions, the researcher remotely observed experiments from a second room.

One computer was setup to capture bio-signal data from the BioRadio device, and was running the BioCapture software (GLNeuroTechnologies, 2015). The second computer was setup to capture video data from the Avermedia Live Gamer Portable device (Avermedia, 2015) and the Microsoft LifeCam HD-3000 web camera, and was running the Open Broadcast Software package (OBSProject, 2015) to facilitate synchronized and overlaid capture of both video sources simultaneously. Additionally, both computers were setup with the Remote Desktop software to allow remote monitoring of the video recording equipment.

### 3.7.3.3. Video Game Setup

Prior to experimental sessions, the Nintendo Wii was turned on, the batteries in the Wii controllers checked, and Tetris Party Deluxe started in preparation. The Live Gamer Portable device was tested to ensure that the game video was both being captured by the recording computer and passed through to the TV.

Tetris Party Deluxe was set to the Marathon Competitive game mode and setup for two players, with bots and items disabled. The TV screen was then turned off until it was time for the experiments to begin.

### 3.7.3.4. Experimental Procedure

Participants were provided with an introductory overview of the research project, including being informed what game they were to play. Participants were shown the BioRadio device, and were informed that it would require the placement of passive electrodes upon their hands and face, and that these sensors cannot cause them any discomfort or harm. The general procedure of the experiment was explained to participants. They were then asked to read the ethics information sheet, and then sign the ethics consent form if they agreed to participate within the experiment.

The procedures used within this study were informed by the results of previous studies utilising psychophysiology (Nacke *et al.*, 2006; Nacke *et al.*, 2009; Nacke *et al.*, 2011; Mandryk *et al.*, 2006; Mandryk *et al.*, 2007; Kivikangas *et al.*, 2011), which include recommendations that researchers should take measures to minimise the interaction between participants and researchers, to reduce the influence of these interactions upon the psychophysiological readings. To this end, the experimenter was not present within the testing environment during game play and instead monitored the room and equipment remotely.

Prior to the commencement of experimentation, participants were asked to fill in a demographic survey, which included details such as their age, gender, occupation,

and hobbies. This survey also asked participants to rate their level of experience with video games, both generally and with Tetris specifically.

The BioRadio sensors were then placed upon participants, and the signal strength tested and calibrated to ensure that electrodes are placed correctly. The placement of electrodes required approximately 15 minutes to complete. A pre-game baseline was then recorded for 3 minutes, during which time participants were asked to watch a video of waves in the ocean, without talking. The purpose of this was to obtain the average physiological state at rest for each participant, as this varies between individuals. After the baselines were recorded, participants were provided with instructions on how to play the game, and were allowed 5 minutes to practice playing the game. This practice was intended to reduce the influence of previous experience upon the affective responses produced throughout the game play, to minimise training effects. No recordings were made during this practice play, and participants were again asked to watch the video of waves in the ocean for 3 minutes prior to the commencement of the experimental conditions.

Participants then began the experimental phase of the procedure. As discussed in Section 3.7.1., there are a total of 2 experimental conditions, which participants undertook in a counterbalanced order to help control against fatigue or training effects.

Each experimental condition lasted between 6 and 8 minutes, depending upon the performance of participants within the game. Immediately after each experimental condition, participants were asked to complete the Self-Assessment Manikin (SAM) survey, Game Experience Questionnaire (GEQ) survey, and the GEQ Social Module as an extension of the GEQ. After the completion of the GEQ survey but prior to the commencement of the next experimental condition, an additional baseline recording was made for 3 minutes under the same conditions as previously described. Participants underwent each experimental condition a total of 4 times, playing the game twice with Equipment On, and twice with Equipment Off. The

psychophysiological signals were removed from one participant and placed upon the other at the halfway point, allowing for data to be collected from both players within the dyad.

At the completion of the experimental conditions, participants were thanked for their time, and the recording instruments and electrodes were removed from their bodies. Participants were fully debriefed as to the purpose of the experiment, and any questions answered. Due to the number of experimental conditions, the total time to completion for each set of participants was between 70 and 90 minutes, depending upon how long participants took to play the game each time.

### 3.7.4. Data Processing

Prior to cleaning and analysis, all heart rate (HR), galvanic skin response (GSR), and electromyography (EMG) data was separated into ten-second periods of time, or epochs, to allow for comparison based on temporal events within the gameplay (Cacioppo, Tassinary and Berntson, 2007; Cowley *et al.*, 2016). This resulted in an average of 168 epochs for each physiological measure, and an average of 672 epochs total per participant. The number of epochs varies by participant due to the differing lengths of time that participants require to complete each game session. The specific temporal game events selected and analysed are presented in subsequent sections.

Several extraneous factors, such as limb movement and talking, are able to influence or interfere with GSR and EMG recordings. As these are recorded through the measurement of minute electrical changes across the body, the most prominent extraneous factor was body movements, which caused increased voltage recordings for the duration of bodily movements. As such, it was required that each epoch was visually scored for movement artefacts prior to analysis. Any epoch that contained more than 4 seconds of movement artefacts were removed from the study and recorded as missing data. All data from all participants was divided and examined in this fashion.

The baseline measure for each physiological data source was normalised according to the specific characteristics of each signal, including de-trending additive physiological signals including Galvanic Skin Response (that tend to increase over time) using rolling baselines and best practice data normalisation practices (Cacioppo, Tassinary and Berntson, 2007; Cowley *et al.*, 2016; Nacke, 2015).

The specific techniques required to clean and process physiological data differ based on the data type. Therefore, the specific techniques used to clean and process HR, GSR, and EMG recordings are discussed in the following sections.

### **3.7.4.1. Heart Rate (HR) Processing**

The HR data was cleaned and processed using HRVAS (Ramshur, 2010), an open source MATLAB toolbox that has been successfully used within several previous studies examining HR data (Ramshur, 2010). The HR data was examined in terms of instantaneous heart rate (HR), which can indicate basic physiological arousal as well as providing limited insights into affective valence (Cacioppo, Tassinary and Berntson, 2007).

While HR measures are typically less prone to movement artefacts, as the photoplethysmography (PPG) HR sensor was placed upon a fingertip, it can be susceptible to some more major movement artefacts, particularly where the sensor has significant pressure applied or where the sensor was pulled away from the skin. This required visual examination of the HR data in ten-second epochs, with segments of data containing more than 4 seconds of erroneous data or artefacts being removed from the study and coded as missing data.

For each epoch, the HR signals recorded through the PPG sensor were used to estimate instantaneous HR at each point in time, which was then averaged across for each 10-second epoch. Additionally, the RR Intervals were corrected for artefacts at a 20% cut off, meaning that every RR interval that differed by more than 20% from the previous was discarded. The HR data was then normalised into a percentage

between 0% and 100% of the maximum global HR scores observed within the participant population.

The normalised HR data was retained as a time series, allowing for the temporal examination of heart rate responses or changes against in-game events or stimuli. At this point, the processed heart rate data was considered ready for analysis.

### **3.7.4.2. Galvanic Skin Response (GSR) Processing**

The GSR data was cleaned and processed using LEDALAB (Benedek and Kaernbach, 2010), an open source MATLAB toolbox that has been used successfully within a number of previous studies examining GSR data (Benedek and Kaernbach, 2010; Plazak, 2016).

Before the any data normalisation processes were conducted, the data was filtered by a low pass filter with a corner frequency of 5Hz to assist in the exclusion of high frequency noise and interference, such as movement artefacts and mains power within the room (Cacioppo, Tassinari and Berntson, 2007; Cowley *et al.*, 2016; Klarkowski *et al.*, 2016). The GSR data was then split into ten-second epochs, allowing for visual inspection and removal of data that was overly contaminated with artefacts. Any epochs containing more than 4 seconds of artefacts were discarded and recorded as missing data.

Following procedures set forth by existing research using GSR (Cacioppo, Tassinari and Berntson, 2007), discussed in Section 3.4.4, due to the natural variations in tonic SCL between individuals, and due to the constant fluctuations in SCL that individuals experience over time, a 3-minute baseline SCL was recorded prior to each experimental condition. The mean of this baseline SCL value was then subtracted from the mean value of each epoch recorded during that experimental condition, resulting in epochs containing GSR values that should have originated predominately from phasic SCRs.



As with HR data, the GSR data was then normalised into a percentage between 0% and 100% of the maximum global scores observed within the participant population. At this point, the processed GSR data was considered ready for analysis.

### 3.7.4.3. Electromyography (EMG) Processing

The EMG data was cleaned and processed using EMGLAB (McGill, Lateva and Marateb, 2005), an open source MATLAB toolbox that has been used successfully within several previous studies examining EMG data.

When measuring emotional responses, facial EMG has been repeatedly demonstrated as a valuable measure of emotional valence (Boxtel, 2010). As previously discussed, this project selected two facial muscles as measurement points for EMG recording, namely the corrugator supercilli (CS) as a measure of negatively valenced emotions and the zygomaticus major (ZM) as a measure of positively valenced emotions (Boxtel, 2010). Facial EMG relies on the measurement of electrical currents generated by the activation of facial muscles, through use of small bipolar electrodes of less than 4mm diameter placed over the muscle locations (Boxtel, 2010).

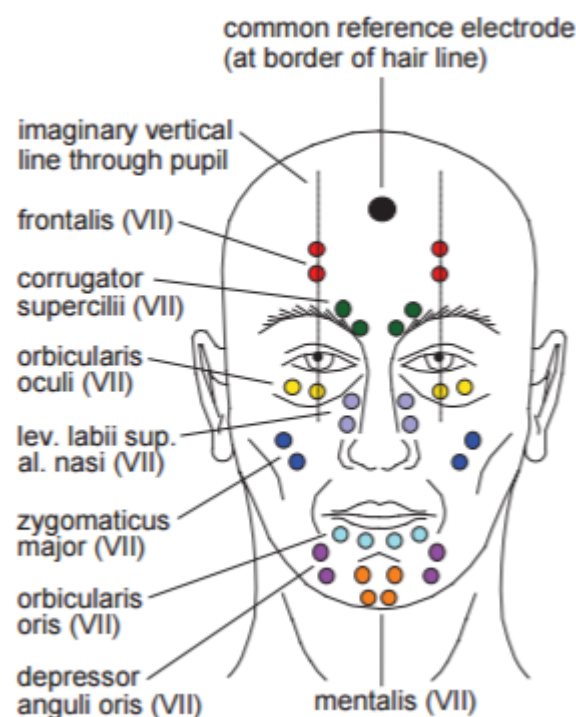


Figure 38: Facial EMG muscle locations and associated electrode placements (Boxtel, 2010).

As with other bio-signals, EMG signals must be carefully filtered and cleaned prior to data analysis to remove signal noise and artefacts. The EMG signal for each muscle was first bandpass filtered between 20-500Hz, as these are the predominate frequency ranges found within facial EMG (Boxtel, 2010). A high-pass filter was then applied at 20Hz, as low frequency artefacts originating from body movements, eye movements, breathing, and swallowing strongly influence the signal at this point and can obscure the desired EMG signals (Cacioppo, Tassinary and Berntson, 2007). Additionally, due to the experiment being conducted outside of an electrically-shielded laboratory, a 60Hz notch filter was applied to remove interference originating from mains power (Cacioppo, Tassinary and Berntson, 2007; Boxtel, 2010).

As with previous physiological measures, the EMG data was then split into 10-second epochs, allowing for visual inspection and removal of data that contained excessive levels of contamination. Any epochs containing more than 4 seconds of artefacts were discarded and recorded as missing data. Like other bio-signals, baseline EMG levels and the amplitude of EMG responses to stimuli vary greatly between individuals, due to physiological differences as well as differences in cognitive affective processes (Cacioppo, Tassinary and Berntson, 2007; Boxtel, 2010; Cowley *et al.*, 2016).

EMG data was normalised as a percentage of the global maximum score (Mandryk, Atkins and Inkpen, 2006; Mandryk and Atkins, 2007). This resulted in the expression of EMG responses as a percentage of maximum response rather than as the value of the difference between baseline and response. This has the additional advantage of allowing for simpler comparisons between muscle activations, between individuals, between experimental conditions, and assists in improving retest reliability for the same individuals at later dates (Mandryk, Atkins and Inkpen, 2006; Boxtel, 2010).

The proportional variances between baseline and affective responses were averaged for both the CS and ZM muscles within each epoch, reducing the raw facial EMG data for each epoch down to four data sources:

1. **Mean proportional CS activation:** A single value per epoch, indicating level of event related negatively valenced emotions.
2. **Mean proportional ZM activation:** A single value per epoch, indicating level of event related positively valenced emotions.
3. **Proportional CS time series:** Allowing for temporal examination of negatively valenced emotions against in-game events.
4. **Proportional ZM time series:** Allowing for temporal examination of positively valenced emotions against in-game events.

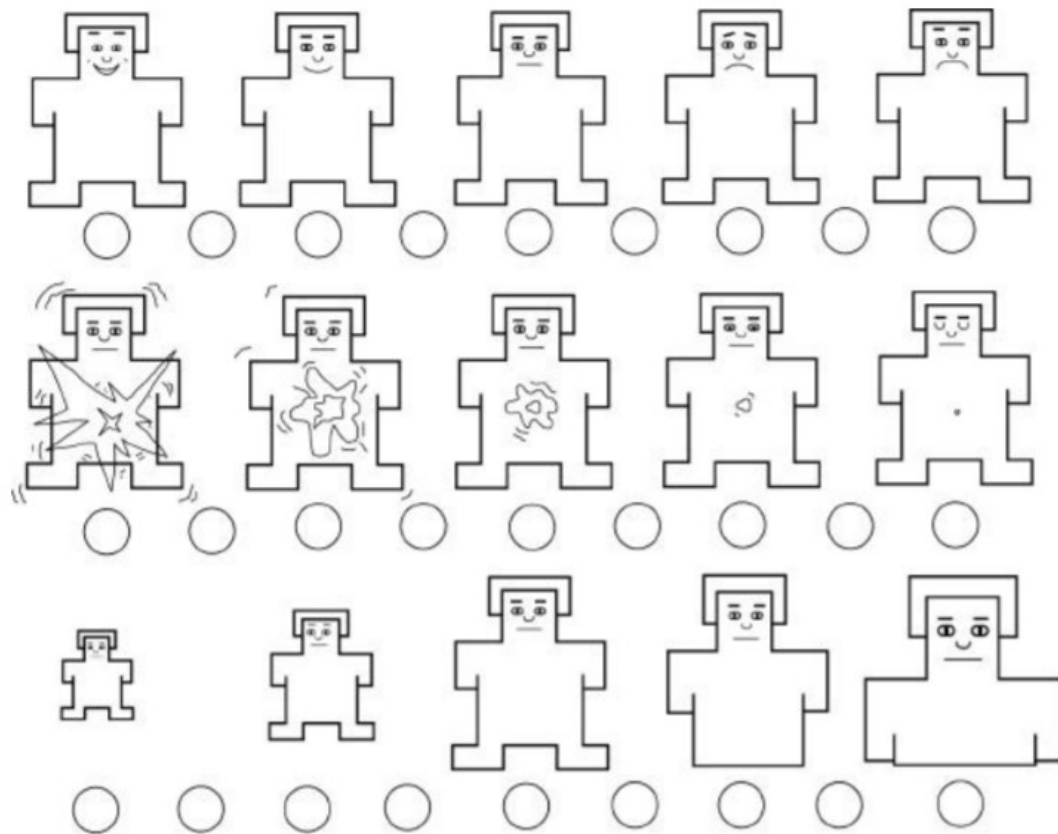
All data contained within the EMG time series was visually inspected from video recordings to assess the accuracy of the EMG measures, with measures appearing to stem from false positives or other events being discarded. At this point, the processed facial EMG data was considered ready for analysis.

#### 3.7.4.4. Self-Assessment Manikin (SAM) Survey Processing

The Self-Assessment Manikin (SAM) is a simple but statistically validated self-report survey that has been used successfully within several multimedia fields, including video game research (Ravaja *et al.*, 2004; Ivory and Kalyanaraman, 2007; Klarkowski *et al.*, 2016). As illustrated by Figure 39, the standard SAM instrument contains 21 images arrayed into 3 rows, each of which represents a dimension of emotion from the Pleasure Arousal Dominance (PAD) model of emotion first proposed by Russell and Mehrabian (1977).

Although the SAM collects three factors relevant to affective responses, namely Pleasure, Arousal, and Dominance, only two are strictly necessary within this project. The measure of Pleasure correlates directly to Valence, which is used in conjunction with Arousal within this study for modelling and comparing affective responses. Dominance is an additional factor that can be used within analysis, however it does

not strictly fit within Russell's circumplex model of emotion (1989) that this project uses for the assessment of affective responses, and so is not necessary for this purpose.



*Figure 39: The Self-Assessment Manikin (SAM) from the IAPS norms (Lang, Greenwald, and Bradley, 1989). The rows (from top to bottom) require individuals to rate their perceived pleasure, arousal, and dominance.*

Processing SAM surveys is relatively simple, with only three numbers being produced per participant. The first row of the SAM tool represents the Pleasure scale, the second row represents the Arousal scale, and the third row represents the Dominance scale (Morris, 1995). All rows provide 9 circles for participants to select between, depending on where they feel they currently fit upon the scale. For each row, the 5<sup>th</sup> (middle) circle represented a neutral (0) score, with every circle to the right adding 1, and every circle to the left subtracting 1. Therefore, the far-left circle was scored as -4, and the far-right circle was scored as +4.

The processed SAM survey therefore comes out quite simply as three numbers, each of which ranges between -4 and +4. For example, if a participant selected the 8<sup>th</sup> circle for Pleasure, the 6<sup>th</sup> circle for Arousal, and the 5<sup>th</sup> circle for Dominance, the outcome would simply be “P = 3, A = 1, D = 0”. This outcome is indicative of a highly pleasurable emotion that is only slightly arousing, such as a general feeling of enjoyment.

To facilitate the comparison between SAM scores and objective physiological scores, SAM scores were adjusted into a percentage of their maximum using the formula illustrated in Figure 40 below. It should be noted that this is not the typical approach to scoring the SAM instrument and may limit the applicability of SAM results to studies of other researchers, however it facilitates better comparisons across the measures for the purposes of addressing the core research objectives of this project.

$$\left( \frac{score + 4}{8} \right) * 100$$

Figure 40: Formula used to normalise SAM results into percentages of maximum.

This normalised the SAM results so that all data was positive, with 50% being neutral, making it assume the same data format as physiological measures of arousal and valence.

### **3.7.4.5. Game Experience Questionnaire (GEQ) Survey Processing**

The Game Experience Questionnaire (GEQ) by Ijsselstein, de Kort and Poels (2008) is one of the most prominent surveys seeking to measure aspects of perceived video game experiences. The GEQ contains multiple components that can be used either individually or in combination, including the two used within this project, namely the GEQ Core Module for eliciting information on the core game experience, and the

GEQ Social Module for eliciting information about social aspects of game experiences.

The GEQ Core Module breaks game experiences down into seven components, while the GEQ Social Module consists of three additional components relating to experiences of social game play. Both requires individuals to provide responses on a 5-point Likert scale between 0 for “not at all” and 4 for “extremely”, and were scored by allocating points (between 0-4) from each question to its respective component.

A total of six GEQ survey modules were completed by each participant: one GEQ Core module per experimental condition (four total), and one GEQ Social module per social experimental condition (two total). Each survey was first checked to ensure that all questions were answered correctly, and then the scores for each component were summed for each experimental condition.

The seven components of the GEQ Core Module, and their respective questions are:

1. Competence: Items 2, 10, 15, 17, 21.
2. Sensory and Imaginative Immersion: Items 3, 12, 18, 19, 27, 30.
3. Flow: Items 5, 13, 25, 28, 31.
4. Tension/Annoyance: Items 22, 24, 29.
5. Challenge: Items 11, 23, 26, 32, 33.
6. Negative Affect: Items 7, 8, 9, 16.
7. Positive Affect: Items 1, 4, 6, 14, 20.

The three components of the GEQ Social Module, and their respective questions are:

7. Psychological Involvement – Empathy: Items 1, 4, 8, 9, 10, 13.
7. Psychological Involvement – Negative Feelings: Items 7, 11, 12, 16, 17.
7. Behavioural Involvement: Items 2, 3, 5, 6, 14, 15.

These scores were summed, and then normalised by dividing the score of each component by the number of questions allocated to each, so that the resulting scores for each component were directly comparable, with scores ranging between 0 and 4 for each. As with the SAM results, these were then converted into a normalised format using percentage of maximums, using the formula below. It should be noted that this is not the typical approach to scoring the GEQ instrument and may limit the applicability of GEQ results to studies of other researchers, however it facilitates better comparisons across the measures for the purposes of addressing the core research objectives of this project.

$$\left(\frac{\text{score}}{4}\right) * 100$$

Figure 41: Formula used to normalise GEQ results into percentages of maximum.

This reduced the GEQ survey data down to 10 measures per experimental condition:

1. **Competence:** Level of proficiency and skill.
2. **Sensory and Imaginative Immersion:** Level of emotional involvement.
3. **Flow:** Level of engagement and focus.
4. **Tension/Annoyance:** Level of mental or emotional strain.
5. **Challenge:** Level of difficulty of the game.
6. **Negative Affect:** Level of negative emotions experienced.
7. **Positive Affect:** Level of positive emotions experienced.
8. **Involvement – Empathy:** Level of empathy for other players.
9. **Involvement – Negative Feelings:** Level of negative emotions for others.
10. **Behavioural Involvement:** Level of interaction between players.

At this point, the processed GEQ survey data was considered ready for analysis.

### 3.7.4.6. Video Recording Processing

The videos recorded during experimental sessions, which captured gameplay footage alongside participant facial expressions. Prior to analysis, Windows Movie

Maker was used to edit each video to remove all excess video recorded before experimental conditions commenced, recorded in between experimental conditions, and recorded after experimental conditions. Videos were split at each condition, so that each experimental condition for each participant was stored as a separate video. This served the purpose of reducing data storage requirements, as well as simplifying data processing and analysis.

Each video recorded was manually examined, and the precise time of events of interest were recorded within the Catalogue of Events (see Section 3.9) for data analysis purposes. Events of interest included evidence of facial expressions, evidence of social interactions between player dyads, and game mechanics or events. The procedures used to code facial expressions, social interactions, and game mechanics are described in Section 3.9.

### **3.7.5. Methodology: Primary Study Summary**

The primary study within this research project aimed to collect data on the affective responses as elicited through interaction with video game events, and additionally to examine measurement effects produced as a result of psychophysiological equipment.

The methodology section for the primary study has presented the design of the method used, including detailing the varying processes and procedures used throughout. The following section discusses how the data collected through this method was modelled into scores of emotion for further analysis.

## **3.8. Modelling of Emotions from Physiological Sources**

Deriving or classifying emotions from physiological data is not a simple process, especially when considering continuous measures, where there is a lack of clear or established guidelines that can be considered to be best practice (Mandryk, Atkins and Inkpen, 2006; Cacioppo, Tassinary and Berntson, 2007; Mandryk and Atkins,



2007; Cowley *et al.*, 2016). This is due to the nature of physiological signals, which are influenced by both sympathetic and parasympathetic nervous activity simultaneously (Cacioppo, Tassinary and Berntson, 2007; Cowley *et al.*, 2016), as well as the rapid development of new technologies that frequently change methods for both physiological measurement and emotional assessment (Cacioppo, Tassinary and Berntson, 2007).

Studies using physiological measures for assessment of affective responses are dominated by the inclusion of GSR, HR, Facial EMG, or combinations (Ravaja and Kivikangas, 2008b; Kivikangas *et al.*, 2011; Jarvela *et al.*, 2014; Cowley *et al.*, 2016), however these are key indicators of arousal and valence rather than direct correlates of emotional responses (Cacioppo, Tassinary and Berntson, 2007). There is a growing body of literature that demonstrates how measures of arousal and valence can be used to infer emotions (Mandryk, Atkins and Inkpen, 2006; Ravaja and Kivikangas, 2008b; Kivikangas *et al.*, 2011; Nogueira, Rodrigues and Oliveira, 2013; B. Cowley *et al.*, 2014), however the approaches used vary considerable, and no single approach reviewed appears to be dominant across domains (see Section 2.3).

Inference of affective responses based upon physiological signals therefore requires an additional step prior to data analysis and interpretation, where the physiological data undergoes modelling to produce classified emotional states based upon current empirical understandings of emotions and their respective bodily influences (Cacioppo, Tassinary and Berntson, 2007; Mandryk and Atkins, 2007; Cowley *et al.*, 2016).

Within the video game and HCI domains, there have been a number of attempts to develop models of emotion based upon psychophysiological measures, facilitating the classification of emotions and affective responses based upon a series of logic rules applied to the physiological data that attempt to account for the multiple factors influencing the signals (e.g. Mandryk and Atkins, 2007; Levillain *et al.*, 2010; Nogueira *et al.*, 2014).

### 3.8.1. Fuzzy Logic Approaches

While it is possible to manually model emotions manually using expert knowledge, doing so requires a great deal of time, and increases the chances of biases being introduced, and does not appear to provide any benefit to the quality of emotion classification (Mandryk and Atkins, 2007).

A range of algorithmic approaches have been attempted for the purposes of modelling emotions, such as neural networks (Fatourehchi *et al.*, 2007; Matiko, Beeby and John, 2014), Naïve Bayes (Koelstra and Muhl, 2012; Matiko, Beeby and John, 2014), k-nearest neighbors (kNN)(Koelstra and Muhl, 2012), and Fuzzy Logic (Mandryk and Atkins, 2007; Blewitt, 2012; Nogueira *et al.*, 2014). Fuzzy Logic is the dominant approach within HCI and video game studies seeking to model emotions and affective responses, and has been used successfully and validated in a range of previous studies (E.g. Mandryk and Atkins, 2007; Levillain *et al.*, 2010; Orero *et al.*, 2010; Nogueira *et al.*, 2014).

Fuzzy Logic is an approach to algorithmic assignment of patterns into classes given certain degrees of membership, where the true value may range between completely true and completely false depending upon the relationships between the data (Kumar *et al.*, 2008; Blewitt, 2012). The success of Fuzzy Logic in video game and HCI fields appears largely due to the way in which it classifies objects based on membership rules that may vary or conflict (Mandryk and Atkins, 2007). The objectives of this project are to assess affective responses to video game events, and so while this necessitates a method for classifying emotions, the creation of a new model of emotions is beyond the scope of this project.

The Fuzzy Logic Model of Emotion first used by Mandryk *et al.* (2006) is arguably the most used and validated of the Fuzzy Logic approaches used within studies utilising psychophysiology, and has been replicated and proven useful in a range of studies since first being introduced in 2006 (e.g. Mandryk and Atkins, 2007; Levillain *et al.*,

2010; Orero *et al.*, 2010; Nogueira *et al.*, 2014). Alternative Fuzzy Logic methods include those proposed by Matiko, Beeby and Tudor (2014), where the authors attempt to model EEG data based upon previous research on brain region activation in response to stimuli, or the approach of Orero *et al.* (2010), who attempt to create their own fuzzy decision tree to assess variations in challenge during video game play.

### 3.8.2. The Mandryk Model of Emotion

While alternative approaches such as these demonstrated their ability to obtain interesting and valuable insights into affective responses to video game play, the approach presented by Mandryk *et al.* (2006; 2007) is clearly documented, has been replicated successfully previously. The Mandryk Model of Emotion uses 22 rules within a mamdani fuzzy system to classify four physiological signals, GSR, HR, EMG<sub>frown</sub>, and EMG<sub>smile</sub>, onto a modified version of Russel's circumplex model of emotion (Russell *et al.*, 1989). The resultant arousal and valence scores are mapped onto a 6x6 grid, referred to as the Affect Grid (see Figure 42).

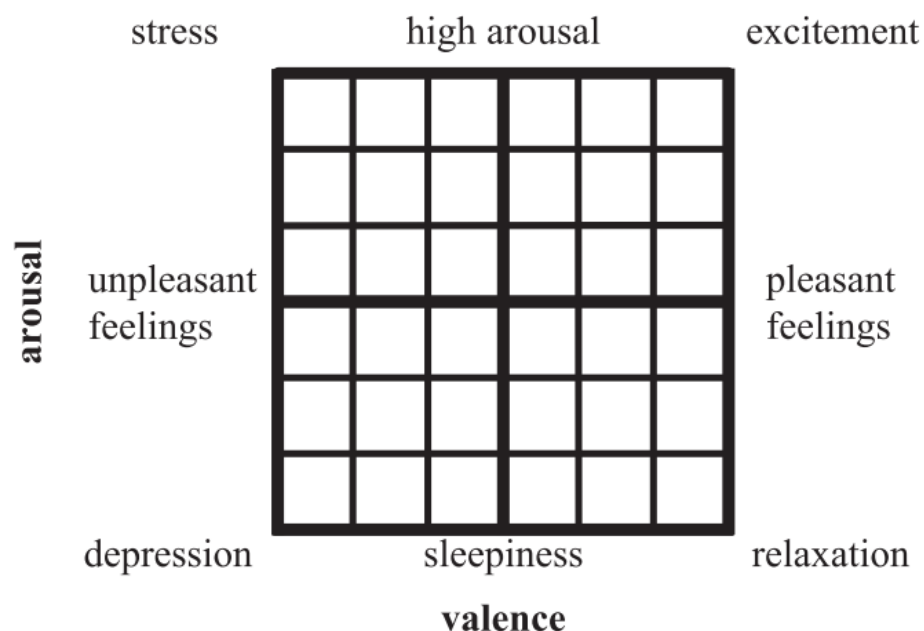


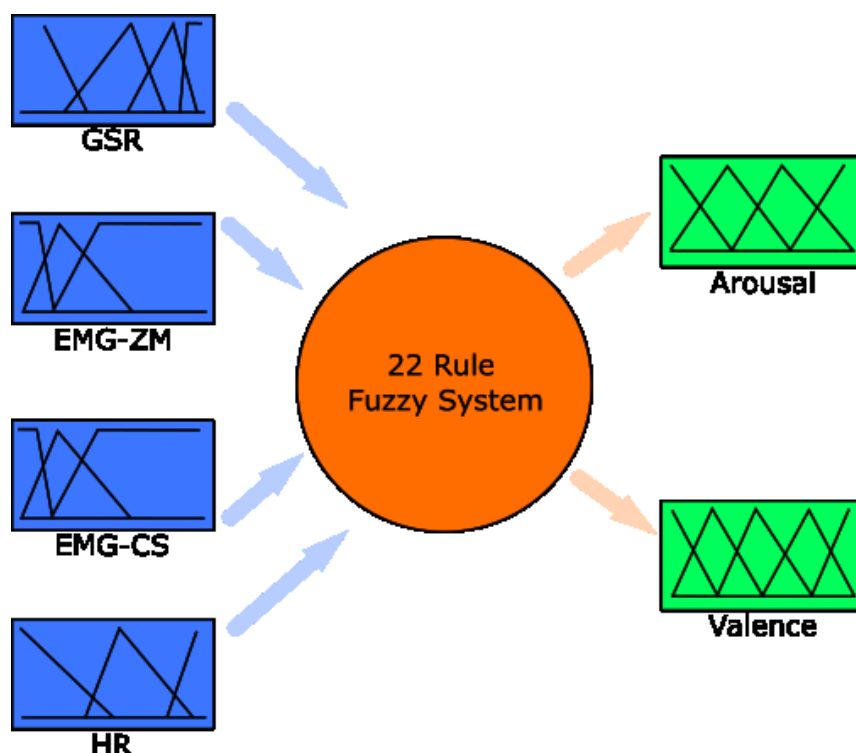
Figure 42: The modified Affect Grid used by Mandryk *et al.* (2006) to model physiological signals and interpret arousal and valence.

It then uses 67 additional rules within a second mamdani fuzzy system to classify the scores on the Affect Grid into five categories of emotions, namely Fun, Challenge, Boredom, Frustration, and Excitement, each with varying strengths. The five categories of emotions and the level of differentiation between them was deemed most appropriate for this project, and so was selected for use.

### 3.8.2.1. Mapping Arousal and Valence (AV space)

The complete time series each physiological signal for each participant are first classified within the AV space, resulting in the generation of a new time series of the arousal and valence for each participant at each point in time (Mandryk and Atkins, 2007).

The normalised physiological data (discussed in Section 3.7.4) for each participant were processed through a mamdani fuzzy system with 4 inputs (GSR,  $EMG_{smile}$ ,  $EMG_{frown}$ , and HR), 2 outputs (arousal and valence), and 22 rules to determine membership functions (detailed in Appendix D).



*Figure 43: The mamdani fuzzy system used by Mandryk et al. (2006) to classify four physiological signals into arousal and valence.*

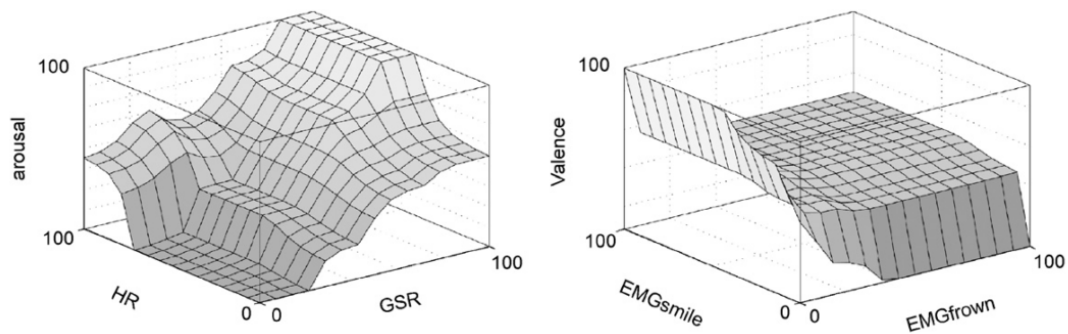
The development of the 22 membership rules is discussed by Mandryk and Atkins (2007) and a critique of them is beyond the scope of this project, so they are briefly summarised below. Membership rules for physiological signals were generated based upon the characteristics of each physiological source, in an attempt to control for variations between physiological measures caused by ANS activity (Mandryk and Atkins, 2007).

Heart Rate (HR) data was classified based upon its respective mean and standard deviation, and are described as low, medium, and high HR activity (Mandryk *et al.*, 2006). Galvanic Skin Response (GSR) data was classified based on its having multiple peaks or areas of activation, and are described as low, mid-low, mid-high, and high GSR activity (Mandryk *et al.*, 2006). Electromyography (EMG) was low for the recording of both smiles and frowns, with a data distribution that was lognormal in nature, and so was described as low, medium, and high EMG activity (Mandryk *et al.*, 2006). For each physiological signal, the data was split into the respective categorical levels based upon relative values.

The resultant data was classified on a time series with four levels of membership for arousal (low, mid-low, mid-high, and high) and with five levels of membership for valence (very low, low, neutral, high, and very high) (Mandryk *et al.*, 2006). All inputs and outputs are entered as percentages of their maximum possible value, based upon data processing and normalisation procedures (see Section 3.7.4). GSR and HR data was used to generate membership for arousal, while EMG<sub>frown</sub>, EMG<sub>smile</sub>, and HR were used to generate membership for valence (Mandryk *et al.*, 2006).

As illustrated in Figure 44 below, the first fuzzy system maps physiological signals into the AV space using a series of rules that attempt to account for variations between data signals. HR and GSR had a reciprocal yet occasionally conflicting

influence upon arousal, while EMG frowns and smiles were largely neutral in nature, but demonstrated positive and negative valences when participants made facial expressions.



*Figure 44: 3D models depicting how the 22 membership rules map physiological signals into varying degrees of arousal and valence. From Mandryk et al. (2006).*

The mapped AV space for each participant was maintained, to allow for comparisons between experimental conditions and game events with arousal and valence as the unit of analysis. Additionally, the mapped AV space time series for each participant were subjected to a second fuzzy system in an attempt to derive measures of five emotions from the AV space, the process for which is now described.

### 3.8.2.2. Mapping Emotions

To develop the model of emotions experienced during video game play, the AV space time series for each participant underwent transformation through a second fuzzy system, mapping arousal and valence information into membership of five categories of emotions (Mandryk et al., 2006). Following the method demonstrated by Mandryk and co-authors (2006), arousal and valence information derived through the AV space mapping were used as inputs into a 67 rule mamdani fuzzy system (rules listed in Appendix D), resulting in classification into 5 outputs (boredom, challenge, excitement, frustration, and fun).

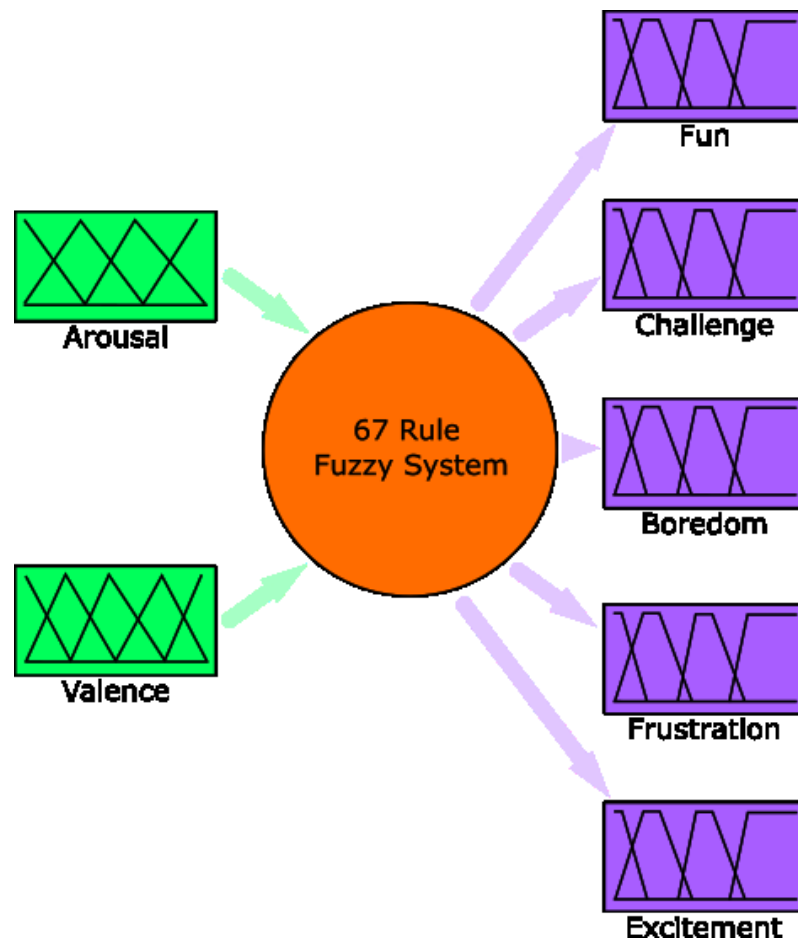


Figure 45: The Emotion Modelling mamdani fuzzy system used by Mandryk *et al.* (2006) to classify arousal and valence (AV space) into five emotional outputs. Each input and output contains a schematic representing the location and form of each respective membership function.

The resultant data was classified on a time series with four levels of membership for each category of emotion (very low, low, medium, and high). All inputs and outputs are entered as percentages of their maximum possible value, based upon data processing and normalisation procedures (see Section 3.7.4). Mandryk *et al.* (2006) state the development of each category of emotion as being based upon guidelines from Russell *et al.* (1989), though note that fun and challenge aren't strictly emotions in the traditional sense. The authors argue for the inclusion of fun and challenge as categories of emotion within their model as they are measurable through arousal and valence, and are of significant value to the research of video games (Mandryk *et al.*, 2006; Mandryk *et al.*, 2007).

As illustrated in Figure 46 below, the second fuzzy system maps the arousal and valence modelled previously onto an Affect Grid based upon the membership rules listed above (Mandryk *et al.*, 2006).

- **Fun** was classified as those affective responses that scored high on valence, but with minimal influence upon arousal.
- **Excitement** was classified as those affective responses that scored high on both arousal and valence.
- **Frustration** was classified as those affective responses that scored high on arousal, but low on valence.
- **Challenge** was classified as those responses that scored highly on arousal, but with relatively neutral valence.
- **Boredom** was classified as those affective responses that scored lowly on both arousal and valence.

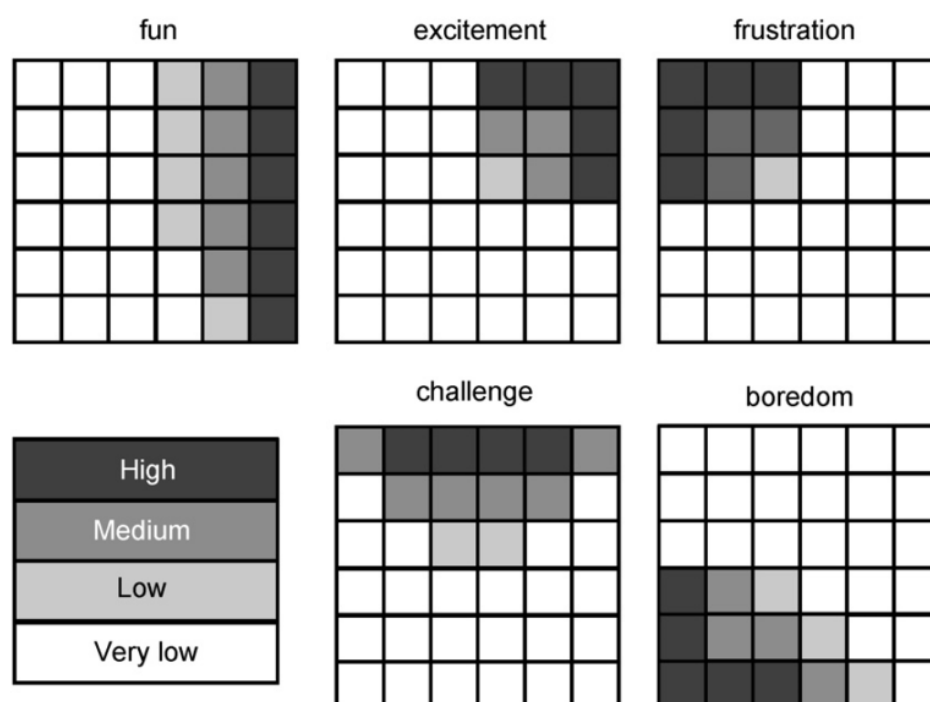


Figure 46: Representation of modelled emotions on the Affect Grid, which classifies emotions based upon affect and valence within the AV space previously discussed. From Mandryk *et al.* (2006).



A model of emotion for each participant was produced as a time series, allowing for comparisons between experimental conditions and game events with emotional classifications as the unit of analysis. The use of a time series facilitates analysis of mean affective responses, averaged across each experimental condition, which can provide insights into the general affective nature of video game play variations in each condition. More importantly, the time series allows for the investigation and analysis from any given point in time, facilitating analysis of specific event related responses or observations.

### 3.8.3. Summary of Fuzzy Model of Emotion

While the processed physiological data shows relative values of bodily responses that are key indicators of arousal and valence in response to stimuli, inferring emotional states and affective responses from physiological data is inherently complicated. This is due to the nature of physiological signals, which are influenced by both sympathetic and parasympathetic nervous activity simultaneously (Cacioppo, Tassinary and Berntson, 2007; Cowley *et al.*, 2016).

Within the fields of video games and HCI research, a number of studies have attempted to model emotional responses through logic algorithms applied to physiological signals, the most dominant approach being the application of fuzzy logic systems. The Mandryk method (Mandryk, Atkins and Inkpen, 2006; Mandryk and Atkins, 2007) for modelling emotions is arguably the approach that is most valid and appropriate for this project, and so its selection and use for the purposes of deriving and modelling the emotional state of each participant during video game play was discussed in the above sections.

Mandryk *et al.* (2006) note that their method is not without limitations, perhaps the most notable being that not all emotions are captured. The implication is that some important emotional responses to video game play are not easily able to be defined within their model. A key example is *shadenfreude*, where individuals relish the

misfortunes of others, which is difficult to classify through arousal and valence as it is characterised through taking pride or gloating when others experience unfortunate events (Mandryk *et al.*, 2006). Improving the classification of emotions is beyond the scope of this project, however in the future additional or improved physiological signals may be able to better detect nuances in responses that are currently difficult or impossible to accurately assess.

The inclusion of additional measures, such as video recording facial responses and social events, is intended to assist in providing additional context on affective responses during video game play and potentially mitigate limitations related to difficulty of classifying some emotions through manual assessment methods. The process through which event related data, including the modelled emotions derived through the fuzzy logic systems described above, is obtained and catalogued is presented within Section 3.9.

### 3.9. Catalogue of Events

#### 3.9.1. Overview

Owing both to the research objectives and the range of data sources used within this project, where the intention was to explore time-sensitive events of interest during video game play, the sum of data available with which to address each research question is unwieldy and requires further reduction. The processed data collected during this project is capable of being used to address the research objectives of this project at both high-temporal resolution, where measures are reported and analysed within the immediate context that each one occurs, and low-temporal resolution, where measures are averaged across each experimental condition or variable. However, generating enriched data that is both accurate and useful for this project therefore requires a further step between the data processing and the data analysis stages, the process for which is discussed below.

Most existing studies using psychophysiology within video game contexts rely upon low-temporal resolution analysis, such as Klarkowski *et al.* (2016) who used mean Galvanic Skin Response (GSR) values as a measure of affective to compare against self-reported affect. This method can therefore provide insights into the average game experience, and the average affective response to video games, flattening out the minutia and simplifying the analysis.

Low-temporal resolution data has the additional benefit of being relatively simple to analyse, as it is essentially a comparison between experimental conditions where the data has been averaged. For example, Klarkowski *et al.* (2016) collect physiological GSR data with the objective of comparing affect between different video game difficulties, process the data to remove artefacts and contaminated data, and then proceed with a repeated-measures ANOVA for the analysis of the averaged affect between experimental conditions. While this is a perfectly valid and appropriate approach given their respective objectives (Cacioppo, Tassinary and Berntson, 2007; Cowley *et al.*, 2016), it comes at the cost of obscuring the events that occur during video game play. This makes it impossible to differentiate between specific time sensitive contexts that contribute towards the overall average affective response, such as the affective responses that might arise from the player being killed, or beating an in-game boss, for example.

In contrast, high-temporal resolution data provides multiple avenues and lenses through which to analyse time sensitive affective responses, allowing much greater insights into affective responses to specific events (Cacioppo, Tassinary and Berntson, 2007; Jones and Rowe, 2015; Cowley *et al.*, 2016), which could theoretically include those that occur during video game play. Within other fields of research, high-temporal event-related studies typically study a single event type multiple times, and average results based on the event (Bartholow, Bushman and Sestir, 2006; Kirmizi-Alsan *et al.*, 2006; Cacioppo, Tassinary and Berntson, 2007). This is useful for the examination of, for example, affective responses to a single image or stimuli in isolation at a known point in time, and has repeatedly been validated for

studies such as the influence of shock stimuli on affective stress or the influence of a violent image versus a happy image on affective valence (Felmingham, Bryant and Gordon, 2003; Olofsson *et al.*, 2008), as just two from the many available examples.

However, this is not a useful approach to studying video games as a stimulus, as even very simple video games are experienced as multiple competing events occurring within close succession. Removing a single video game stimulus and requiring participants to experience it independently may result in furthered understandings of the influence that this specific game mechanic or event has upon affective responses, but it also fundamentally removes this event from the context of video game play, and thereby contaminates anything that might be learned about the affective experience.

While high-temporal resolution data is inherently more complex than low-temporal resolution data, the traditional investigation of only a single stimulus still allows for the data to be processed down to an average temporal response across the participant population for that one affective response (Cacioppo, Tassinary and Berntson, 2007; Hazlett, 2006). However, this project is primarily interested in the investigation of affective responses to events experienced during naturalistic video game play, which entails the experience of multiple competing stimuli at relatively arbitrary or otherwise unknowable points in time. This requires a method of placing multiple stimuli within an event timeline, and analysing relative affective responses to each of those stimuli as they occur, and relative to the contexts within which they occur.

### 3.9.2. Novel Design: Catalogue of Events (CoE)

To mitigate the limitations of both low-temporal and high-temporal resolution methods, a novel method was designed based upon traditional high-temporal resolution event-related psychophysiological methods such as those used by Mandryk *et al.* (2006; 2007), and Ravaja *et al.* (2004; 2006), but modified to facilitate dynamic and continuous video game contexts and environments. To this end, a

Catalogue of Events (CoE) was created, being a form of derived high-temporal resolution analysis of affective responses, and combined four processed physiological measures, coded observations of video game events, coded observations of facial expressions, and coded observations of social interactions into a catalogue that spans all measured and observed events experienced during video game play. As discussed within Section 3.9.3 below, whenever an event was observed within any data source, whether it be a physiological response, a facial expression, a game interaction, or a social interaction, a record was created within the Catalogue of Events including processed forms of all data sources for that point in time. For details of the respective data processing procedures used for each data source prior to events and contextualised data being placed within the Catalogue of Events, refer to Section 3.7.4.

The Catalogue of Events contains a list of discrete observed events that occurred throughout the primary experiment. Each event is logged with a participant identifier, a start time, an end time, a duration in seconds, a coded name for the event, a note for the experimental condition it occurred within, the mean percentages of the maximum value for arousal, valence, and the 5 categories of emotions derived from the fuzzy logic systems (see Section 3.8.2.2), coded facial expressions, coded social interactions, coded game events (see Section 3.9.3), and any relevant annotations.

The catalogue was stored as a Comma Separated Values (CSV) file, an open data standard that is highly supported in a wide range of data analysis software packages, reducing data setup and import time for later data analysis.

The inclusion of affect codes and participant identifiers is intended to facilitate greater depth of analysis, whereby events within the Catalogue of Events can be grouped together by type and analysed by affective response, to assess the influence of unique stimuli upon affective responses; by affective response, to assess the types of events that contribute towards specific affective responses; by relationship to

other event types, to assess the influence that each event type has upon other event types and thereby identify any multiplicative or reductive influences that such conflicts can have; and so forth. It is this level of analytic flexibility that renders this novel method more useful for video game contexts than either traditional low-temporal or high-temporal resolution methods, combining many of the benefits of both and avoiding many of the negative aspects.

### 3.9.3. Registering and Coding Events into the Catalogue

The specific procedures used to detect events of interest from the measures used within this project, and details how they were coded and entered into the Catalogue of Events, are listed in Appendix E. Table 16 summarising these is provided below for ease of reference.

	Detection:	Duration:	Notes:	Registered:
<b>Facial Expressions</b>	Displayed facial expression. Valence recorded	Time between start and end of expression	Intensity coded between 1 and 4	Valence, intensity, and duration of expression
<b>Social Interactions</b>	Interaction between participants	Time between start and end of interaction	Type of interaction recorded	Type, descriptor, duration
<b>Game Events</b>	Mechanics and interactions within the video game	Time between start and end of mechanic	Description of event type and contexts coded	Type, descriptor, contexts, duration
<b>Affective Events</b>	Objective measures of arousal and valence, and modelled emotions	Time between start and end of affective response	Measures modelled using fuzzy logic systems	Arousal, valence, modelled emotions, duration

Table 16: Summary of procedures used to register and enter events into the Catalogue of Events.

### 3.9.4. Event-Related Windows for Physiological Data

For each game event identified, the physiological data was examined on an Event-Related basis. Each event was examined through an Event Window of 10 seconds duration, with 5 seconds being before the event of interest and 5 seconds being after the event of interest.

The Event Window was the same for each physiological signal. This was because the Event-Related data was examined after the processing of physiological data, which included splitting the data into epochs and the normalisation of the respective signals. The normalised and time-corrected physiological signals are then considered fit for direct comparison within the Event Window and Catalogue of Events.

The size of the Event Windows was determined through examination of literature on physiological research into affective responses (Cacioppo, Tassinary and Berntson, 2007; Kivikangas et al., 2011; Cowley et al., 2016). While the 5 seconds prior to the event occurrence is not typical of event-related research (typically studies use less than one second prior to the event), due to the nature of video games and their inherently overlapping mechanics and events, it was considered important to capture events within the context of their respective antecedents.

### 3.9.5. Limitations

While this novel method has a range of benefits and uses that make it a compelling tool for the analysis of affective responses within video game contexts, there are several potential limitations that require discussion.

Firstly, unlike traditional event-related high-temporal methods, this method does not control for cross-stimulus contamination, as it does not measure the video game stimuli or events in isolation. This level of contamination would, in traditional methods, lead to the rejection of the data and the insights derived, as it would not

be possible to form firm conclusions about the affective response to the stimuli of interest (Cacioppo and Tassinary, 1990; Cacioppo, Tassinary and Berntson, 2007; Kivikangas *et al.*, 2011; Cowley *et al.*, 2016). As previously discussed, isolating stimuli is not possible or desirable for studying video game stimuli within contextual environments. However, this also means that confounding influences are not controlled for, and therefore each event recorded may have additional antecedent factors that contribute towards the affective response measured.

It is acknowledged that this may compromise the accuracy of individual events recorded, reducing the precision of the knowledge produced due to video game events not being experienced identically between participants and therefore manipulating the antecedent factors for each recorded affective response (Cacioppo and Tassinary, 1990; Cacioppo, Tassinary and Berntson, 2007). However, the nature of video games is such that even if studies attempted to control for these factors, they would ultimately be unable to do so without compromising the core game experience.

Additionally, video games vary from each other greatly, and even games that are very similar will differ in terms of the mechanics used, the order of presentation, the nuances of game play, the aesthetics, and so forth. This means that understanding specific game mechanics or interactions is not likely to be generalizable beyond vague guidelines such as *“online competitive multiplayer may increase tension and frustration compared to local competitive multiplayer.”* This is therefore not so much a limitation of this novel design, but rather a limitation of understanding video games more generally.

Within this novel method, antecedent factors that may confound affective responses are also recorded within the CoE, and it is possible to analyse each event including relationships with such factors. This can be used to develop additional knowledge on the pathways influencing eventual affective responses during video game play. Additionally, within the context of video games, all game experiences inherently



consist of multiple competing stimuli, and so it would be a logical fallacy to assume that each event could be experienced and therefore understood in isolation from the broader context (Cacioppo, Tassinary and Berntson, 2007; Ravaja et al., 2008).

The second limitation of this novel method is that, compared to traditional methods, it is time intensive, requiring the use of multiple physiological and subjective instruments, and extensive data processing. To draw an example from just one of the data sources, facial videos must be carefully scrutinised, with facial expressions of participants coded and entered within the Catalogue of Events alongside corresponding physiological data. Not only does this require viewing each recorded facial video multiple times to ensure that all facial expressions are captured, it also requires scrutinising all other data sources to collect and code additional data for that event, so that it can be enriched within the Catalogue of Events.

Additionally, while this method facilitates far greater flexibility for analysis of varying influences upon affective responses, it corresponds with a proportional increase in analytical complexity. To draw an example from the Catalogue of Events, the data can be analysed by types of affective responses, by time, by types of game events, and so forth. Each of these units of analysis requires additional analytic effort. Despite this, the depth and flexibility of insights that can be developed through this method go far beyond those that can be developed through traditional methods.

### 3.9.5. Summary

The design of a novel method for the high-temporal resolution analysis of affective responses within naturalistic video game contexts, while departing from traditional techniques and guidelines such as the isolation of stimuli, has proven to be a useful and valuable tool for the analysis of affective responses to dynamic stimuli within the contexts of video game play. Despite limitations on generalisability due to the inclusion of confounding influences upon elicited affective responses, the insights that can be developed through this method can prove highly valuable where the goal

is to understand the events, or sequence of events, experienced during video game play that produce specific affective responses.

### 3.10. Methodology Summary

This methodology chapter has discussed the methodological foundations underpinning this research project, including the identification of the varying methods, techniques, and instruments used in collecting the research data for the purposes of addressing the research questions of this project.

Specifically, the selection of the instruments and measures used within this study have been identified and justified, the experimental procedures for conducting the experiments undertaken within this study have been explicated, and the procedures undertaken for data processing and analysis have been discussed.

The following chapter presents the results of statistical tests conducted upon the data collected, providing insights into the game events identified and how they appeared to influence the elicitation of affective responses.

## Chapter 4: Results

### 4.1. Layout of Chapter / Structure

This chapter provides details on the analytic techniques undertaken with the intention of exploring the data collected through the experimentation conducted, which was described in the previous chapter.

Specifically, statistical results are provided for each of the ten observed game events, alongside and compared against results for overall game play experiences:

- **Mean game play:** Objective measures averaged across play sessions for the aggregate participant population. Events are compared against mean game play to establish differences between event-related responses and general experiences.

1. **Game Speed Increase:** Events where the game speed periodically increased in a cumulative mechanic, resulting in players continually having to respond more quickly.
2. **Non-Useful Tetriminos:** Events where players received Tetriminos that were not useful for their current game board composition, and thereby hindered performance.
3. **Correct Placement:** Events where players placed Tetriminos where they had intended, in a location that either improved performance, or at least did not hinder it.
4. **Incorrect Placement:** Events where players placed Tetriminos that either hindered performance, or where players responded in a way to indicate that the Tetriminos were misplaced.
5. **Multiple Line Clear:** Events where players cleared multiple lines of Tetriminos from the game board simultaneously.
6. **Player Win:** Events where players won the game against their opponent.
7. **Player Loss:** Events where players lost the game against their opponent.
8. **Opponent Correct Placement:** Events where the player observed or commented on their opponent placing Tetriminos in a location that improved their respective performance.
9. **Opponent Incorrect Placement:** Events where the player observed or commented on their opponent placing Tetriminos in a location that hindered their respective performance.
10. **Opponent Multiple Line Clear:** Events where the player observed or commented on their opponent clearing multiple lines of Tetriminos from the game board simultaneously.

Additionally, correlations between subjective and objective measures are provided, as well as results for the influence of physiological measures upon game play experiences. This chapter therefore provides insights into the research questions stated in Chapter 3.

### 4.2. Research Question 1

#### ***Research Question 1:***

*How do events occurring during video game play influence event-related affective responses?*

Research Question 1 aimed to explore the influence that certain video game events can have upon affective responses as experienced during real-time game play. The unit of measurement was therefore observed game events, analysing the influence that these had upon affective responses during game play. This required use of modelled arousal and valence (AV space), as well as the subsequently modelled emotional categories of fun, excitement, boredom, frustration, and challenge (see Section 3.8.2.2 for further details on the modelling process).

The modelled AV space and emotional categories provide real-time objective insights into affective states and responses to game events of interest, which were observed and recorded as part of the data processing prior to analysis (see Section 3.9). The following sections present the results of descriptive and inferential statistical analysis completed for each game event, and the respective influence that each had upon affective responses within this study.

Addressing Research Question 1 requires identifying game events through observation as the independent variable (IV), and collecting objective data relating to the time at which these events occurred as the dependent variables (DVs). The study design is therefore within-participant (repeated-measure), as the DVs are paired to participants, with each participant experiencing multiple events during their game play experience. Affect for the IV game events are compared against the mean affect of the game experience as a whole for each participant (Event Affect vs. Mean Affect), allowing for analysis to determine the effect size, and test whether the mean differences in affect were significant.

Analysing relationships between game events is beyond the scope of this project, and so game events were considered and analysed independently from each other, resulting in a single IV at two levels (observed event, and mean game play baseline) for each test used. As the modelled AV space and emotional categorical data developed through the fuzzy logic data processing are presented in percentages of maximum (see Section 3.8.2), the DVs constitute continuous (ratio) data.

Considering the study design and data types used, the analysis of each game event therefore consists of:

- **Descriptive statistics:** Describing the modelled data for AV and Emotion Types, including each of arousal, valence, fun, excitement, boredom, frustration, and challenge. Additionally, data on observed facial expressions is included to provide further depth.
- **Box plots:** Illustrating the distribution of the data collected, and provides a method for identifying outliers within the dataset.
- **Affect Grid:** A modified circumplex affect grid (see Section 3.8.2.1) illustrating arousal and valence data upon a two-dimensional plane. The mean and standard deviation for each event modelled are included, to illustrate typical player responses.
- **Paired-samples t-tests:** Paired-samples t-tests are commonly used to investigate whether mean differences between paired observations are statistically different from zero (Howell, 2012). Given that RQ1 requires identifying game events and assessing the effect that each has upon affective responses, and with the study design and data types previously discussed, a paired-samples t-test is an appropriate test.

These address Research Question 1 through providing descriptive insights into the effects of specific game events upon affect during video game play, as well as evidence to support the strength and significance of these effects.

Alternative inferential statistical tests considered included one-way MANOVAs, and repeated-measures ANOVAS. In the case of the one-way MANOVA, it was considered inappropriate for the analysis of RQ1 due to the requiring all data to be independently observed, without any within-participant interaction (Howell, 2012). Additionally, it analyses linear composites or vectors of means between participant groups within the independent variable, and therefore requires that the dependent variables correlate within each other in a linear fashion (Howell, 2012). Due to the ways in which the emotional categories are classified using the fuzzy logic technique (see Section 3.8.2.2), the dependent variables do not have linear relationships with one another.

The second best alternative inferential statistical test considered was the repeated-measures ANOVA, which compares changes in a single dependent variable paired to participants at multiple points in time, or after differing conditions. Although the design and data contained within this study meet the requirements of the repeated-measures ANOVA, it is used to compare multiple instances of a single independent variable against a single dependent variable. In the case of this study, this would provide evidence of mean differences between observed game events for a single modelled emotion at a time. For example, it could be used to assess whether modelled Fun differs significantly between Game Speed Increase events and Correct Tetriminos Placement events. However, this level of analysis was not required to answer the research question, and therefore falls under future research.

### **4.2.1. Overall Objective Game Play Experience**

This section provides results for the aggregate, or mean, game play experience as it relates to affective responses at the game level. This required averaging the objective data across the entire game play experience for each participant

individually for later pairwise comparisons, as well as averaging the data across the entire participant population to provide insights into the overall mean affective experience. This serves as a baseline for comparing game events against, in order to determine effect sizes and significance of discrete game events and the influence that they have upon the overall affective experience.

#### 4.2.1.1. Arousal and Valence

As illustrated in Table 17 below, when averaged across the entire participant population, the aggregate arousal was 58.77% (+ 14.53%) and the aggregate valence was 52.15% (+ 18.72%). A Shapiro-Wilk test for normality of distribution was conducted for both arousal and valence, both passing with  $p = .588$  and  $p = .862$  respectively (refer to Table 18).

	N Statistic	Mean Statistic	Std. Deviation Statistic
Aggregate Arousal	44	58.77	14.54
Aggregate Valence	44	52.15	18.72

Table 17: Descriptive statistics for aggregate affect and arousal, averaged across the entire game play session for each participant. Data in % of maximum possible score.

#### Mean Modelled Arousal and Valence: Tests of Normality

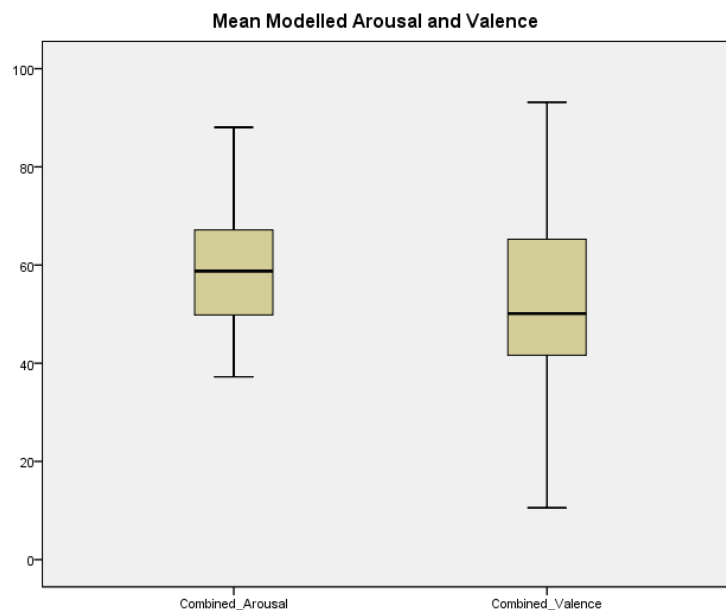
	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Aggregate Arousal	.066	44	.200*	.979	44	.588
Aggregate Valence	.060	44	.200*	.986	44	.862

\*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Table 18: Results of Shapiro-Wilk test for normality, conducted for the aggregate arousal and valence scores. Both arousal and valence were found to be normally distributed.

As illustrated in Figure 47 below, no outliers were evident upon inspection of a box plot for values greater than 1.5 box-lengths from the edge of the box.



*Figure 47: Box plot illustrating the distribution of aggregate data for arousal and valence. Note that no outliers were identified.*

The aggregate modelled arousal and valence scores were mapped onto the Affect Grid, illustrated in Figure 48 below, demonstrating that the overall game experience was mid-level arousing, and with a low-positive valence. The standard deviation for both arousal and valence is represented through the proportionally skewed outer ellipse.



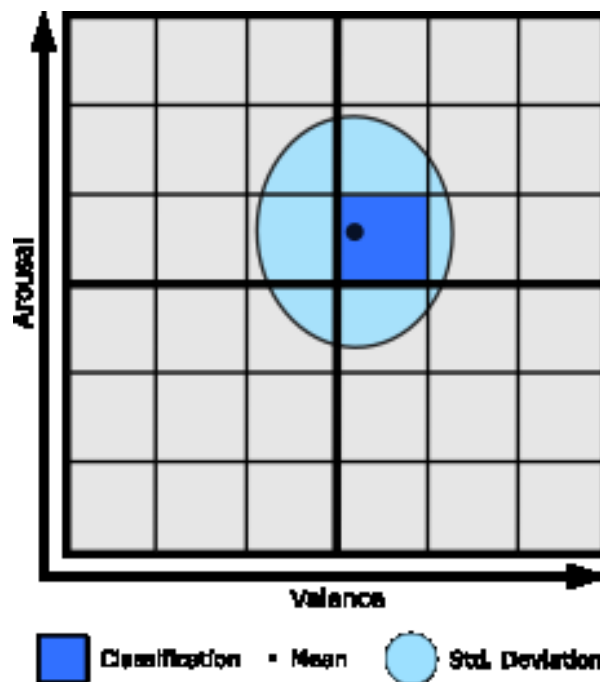


Figure 48: Affect Grid modelling the mean arousal and valence scores, illustrating that the overall game experience was slightly arousing with a positive valence.

#### 4.2.1.1.1. Modelled Emotional Categories

As illustrated in Table 19 below, the aggregate Fun was 61.07% (+/- 11.17%), aggregate Excitement was 55.05% (+/- 13.72%), aggregate Boredom was 23.04% (+/- 8.90%), aggregate Frustration was 21.21% (+/- 8.74%), and aggregate Challenge was 34.45% (+/- 9.19%). A Shapiro-Wilk test for normality of distribution was conducted for all the above, all passing with non-significant *p-values* (refer to Table 20 below).

**Mean Modelled Emotions: Descriptive Statistics**

	N	Mean	Std. Deviation
Aggregate Fun	44	61.07	11.17
Aggregate Excitement	44	55.05	13.72
Aggregate Boredom	44	23.04	8.90
Aggregate Frustration	44	21.21	8.74
Aggregate Challenge	44	34.45	9.19

Table 19: Descriptive statistics for aggregate modelled emotional categories (fun, excitement, boredom, frustration, and challenge). Each of these emotional categories was averaged across the entire game play session for each participant. Data in % of maximum possible score.

**Mean Modelled Emotions: Tests of Normality**

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Combined_Fun	.073	44	.200*	.982	44	.729
Combined_Excitement	.085	44	.200*	.965	44	.195
Combined_Boredom	.095	44	.200*	.974	44	.405
Combined_Frustration	.113	44	.194	.965	44	.208
Combined_Challenge	.099	44	.200*	.980	44	.634

Table 20: Results of Shapiro-Wilks test for normality, conducted for the aggregate fun, excitement, boredom, frustration and challenge scores. All modelled emotions were found to be normally distributed.

As illustrated in Figure 49 below, no outliers were evident upon inspection of a box plot for values greater than 1.5 box lengths from the edge of the box.

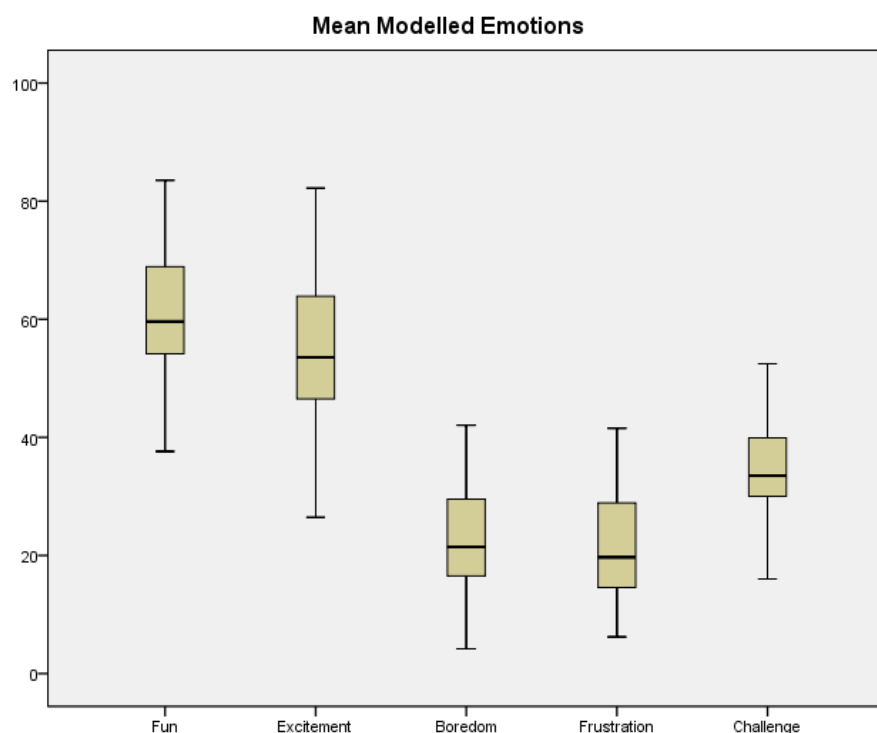


Figure 49: Box plot illustrating the distribution of aggregate data for each modelled emotion. Note that no outliers were identified.

#### 4.2.1.1.2. Observed Facial Expressions

Observed facial expressions are reported in mean duration, in seconds, of either positive or negative expressions being observed (see Section 3.9.3 for further details). As illustrated in Table 21 below, the aggregate duration of positive facial expressions was 1.7 seconds (+/- 0.5 seconds), while the aggregate duration of negative facial expressions was 1.17 seconds (+/- .34 seconds). A Shapiro-Wilk test for normality of distribution was conducted for both positive facial expressions and negative facial expressions, both passing with  $p = .106$  and  $p = .188$  respectively. The data was inspected for outliers through inspection of a box plot, with no outliers identified.

**Observed Facial Expressions: Descriptive Statistics**

	N	Mean	Std. Deviation	Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Std. Error
Combined_PosFacial	44	1.7008	.50472	-.773	.702
Combined_NegFacial	44	1.1714	.34074	2.305	.702

Table 21: Descriptive statistics for aggregate positive and negative facial expressions observed during game play. Data is mean expression duration in seconds.

#### 4.2.2. Comparison of Events with Overall Game Experience

This section provides descriptive results for modelled arousal, valence, and affective responses to ten events observed during game play, separated into four categories of event types (Speed Increase, Tetriminos Placement, Player Performance, and Opponent Performance). Additionally, the data for each of the observed events is compared against the previously discussed overall game play experience, providing insights into the specific influences that each event had upon affective responses during game play.

#### 4.2.2.1. Event Category: Game Speed

Increases of Game Speed (coded as “Game Speed Increase” events) occur as players progress in score during Tetris Party Deluxe. The increased speed causes the Tetriminos to fall more rapidly, reducing the time available for players to react and correctly place them.

##### 4.2.2.1.1. Game Speed Increase event analysis

During game speed increase events, the mean arousal and valence were 77.00% (+19.76%) and 57.67% (+12.49%) respectively. The results for each of the five modelled emotions are illustrated in Table 22 below. A Shapiro-Wilk test for normality was conducted for all data, with all passing and therefore being normally distributed (see Table 23 below).

**Speed Increase Events: Descriptive Statistics**

	N	Mean	Std. Deviation
SpeedIncrease_Arousal	44	77.01	10.76
SpeedIncrease_Valence	44	57.67	12.49
SpeedIncrease_Fun	44	65.67	14.49
SpeedIncrease_Excitement	44	64.44	12.22
SpeedIncrease_Boredom	44	20.27	11.63
SpeedIncrease_Frustration	44	25.53	12.61
SpeedIncrease_Challenge	44	36.45	16.95

Table 22: Descriptive statistics for modelled AV space and emotions, averaged across game event occurrences for each participant. Data in % of maximum possible score.

**Speed Increase Events: Tests of Normality**

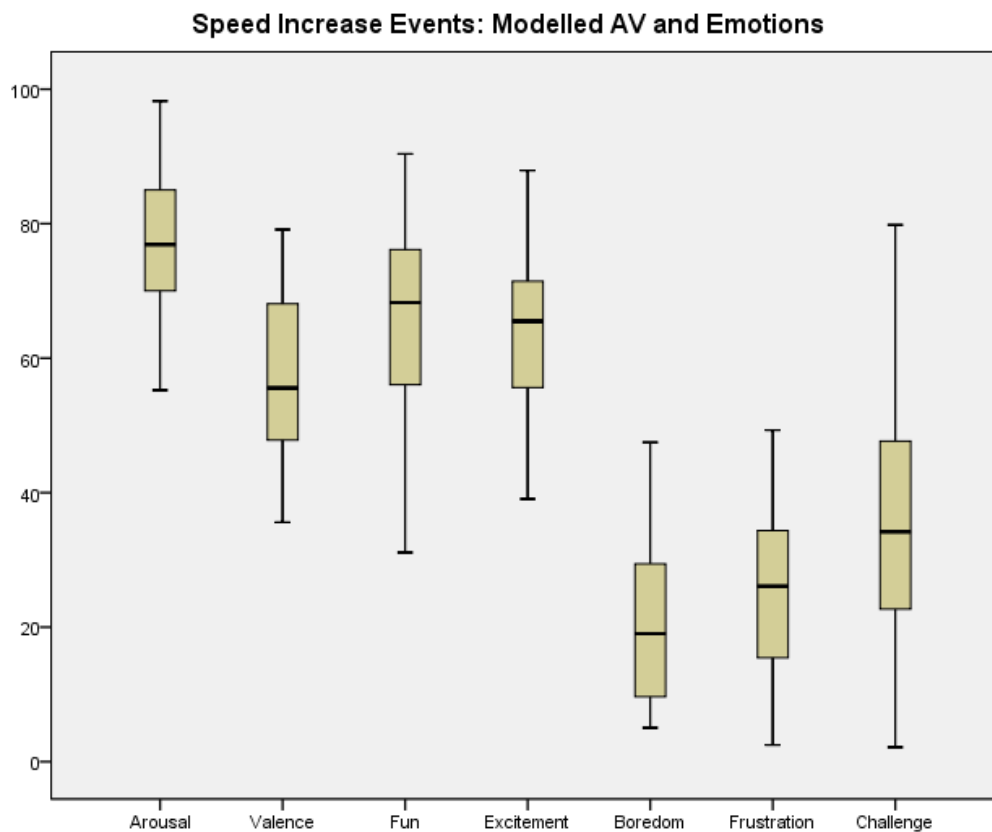
	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
SpeedIncrease_Arousal	.057	44	.200*	.986	44	.866
SpeedIncrease_Valence	.109	44	.200*	.954	44	.176
SpeedIncrease_Fun	.105	44	.200*	.984	44	.774
SpeedIncrease_Excitement	.098	44	.200*	.985	44	.838
SpeedIncrease_Boredom	.167	44	.004	.933	44	.113
SpeedIncrease_Frustration	.058	44	.200*	.976	44	.484
SpeedIncrease_Challenge	.088	44	.200*	.980	44	.647

\*. This is a lower bound of the true significance.

### a. Lilliefors Significance Correction

Table 23: Results of Shapiro-Wilk test for normality, conducted for modelled AV space and emotions. All data was found to be normally distributed.

As illustrated in Figure 50 below, no outliers were evident upon inspection of a box plot for values greater than 1.5 box-lengths from the edge of the box.



*Figure 50: Box plot illustrating the distribution of AV space and emotions occurring during Speed Increase events. Note that no outliers are identified.*

The modelled arousal and valence scores for Speed Increase events were mapped onto the Affect Grid, illustrated in Figure 51 below, demonstrating that participant arousal was mid-level and valence was low-positive. The standard deviation for both arousal and valence is represented through the proportionally skewed outer ellipse.

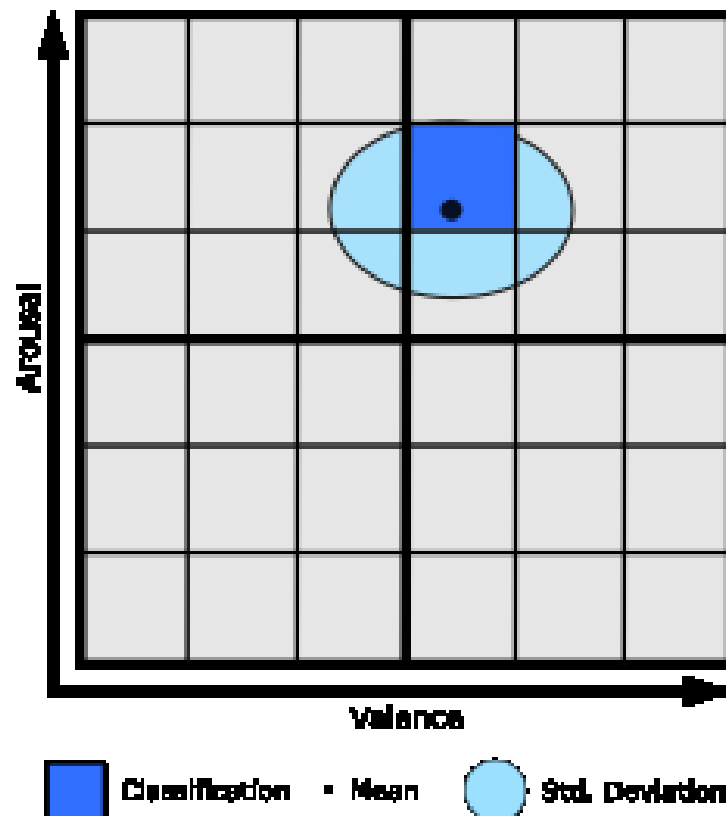


Figure 51: Affect Grid modelling the mean arousal and valence scores occurring during Speed Increase events, illustrating that the affective responses were moderately arousing.

Paired-sample t-tests comparing means for modelled emotions against a baseline of the overall game play experience revealed that Speed Increase events elicited a range of influences upon modelled emotions, as illustrated in Table 24 below:

- **Similar Fun:** A small strength (*Cohen's d* = .22) statistically *insignificant* mean increase of 4.6 (95% CI, -1.79 to 10.99) in modelled fun compared to the overall game experience,  $t(43) = 1.45$ ,  $p = .153$ .
- **Increased Excitement:** A medium strength (*Cohen's d* = .46) statistically *significant* mean increase of 9.16 (95% CI, 3.12 to 15.19) in modelled excitement compared to the overall game experience,  $t(43) = 3.06$ ,  $p = .004$ .
- **Similar Boredom:** A small strength (*Cohen's d* = .21) statistically *insignificant* mean decrease of -2.7 (95% CI, -6.84 to 1.29) in modelled boredom compared to the overall game experience,  $t(43) = -1.376$ ,  $p = .176$ .

- **Increased Frustration:** A small strength (*Cohen's d* = .30) statistically **significant** mean increase of 4.32 (95% CI, -.12 to 8.76) in modelled frustration compared to the overall game experience,  $t(43) = 1.96, p = .041$ .
- **Similar Challenge:** A small strength (*Cohen's d* = .17) statistically **insignificant** mean increase of 1.88 (95% CI, -3.68 to 7.44) in modelled challenge compared to the overall game experience,  $t(43) = .68, p = .499$ .

### Paired Samples Test: Speed Increase Events vs. Mean Game Play

		Paired Differences					t	df	Sig. (2- tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	SpeedIncrease_Fun - Combined_Fun	4.60	21.01	3.17	-1.79	10.99	1.45	43	.153
Pair 2	SpeedIncrease_Excitement - Combined_Excitement	9.16	19.85	2.99	3.12	15.19	3.06	43	<b>.004</b>
Pair 3	SpeedIncrease_Boredom - Combined_Boredom	-2.77	13.38	2.02	-6.84	1.29	-1.37	43	.176
Pair 4	SpeedIncrease_Frustration - Combined_Frustration	4.32	14.59	2.20	-.12	8.76	1.96	43	<b>.041</b>
Pair 5	SpeedIncrease_Challenge - Combined_Challenge	1.88	18.41	2.76	-3.68	7.44	.68	43	.499

Table 24: Results of paired-sample t-tests comparing mean Speed Increase events against mean overall game experience. Mean differences and the significance of these are demonstrated for each of the five modelled emotions.

#### 4.2.2.2. Event Category: Tetriminos Placement

Events that relate to the placement of Tetriminos (coded as “Non-Useful Tetriminos”, “Correct Placement”, and “Incorrect Placement”) occur as players receive and attempt to place Tetriminos within the game board. Receipt of non-useful Tetriminos can hinder the ability of players to correctly place them upon the board, potentially inhibiting performance or resulting in mistakes being made. Likewise, placing Tetriminos either correctly or incorrectly has consequences for player choice and performance moving forward within the game.

#### 4.2.2.2.1. Non-Useful Tetriminos Received

During non-useful Tetriminos received game events, the mean arousal and valence were 63.39% (+ 17.66%) and 48.00% (+ 11.54%) respectively. The results for each of the five modelled emotions are illustrated in Table 25 below. A Shapiro-Wilk test for normality was conducted for all data, with all passing and therefore being normally distributed (see Table 26 below).

**Non-Useful Tetriminos Received Events: Descriptive Statistics**

	N	Mean	Std. Deviation
NonUseful_Valence	44	48.01	11.54
NonUseful_Fun	44	53.08	9.66
NonUseful_Excitement	44	58.27	10.59
NonUseful_Boredom	44	23.41	9.67
NonUseful_Frustration	44	38.18	12.54
NonUseful_Challenge	44	33.91	12.87

*Table 25: Descriptive statistics for modelled AV space and emotions, averaged across each time non-useful Tetriminos were received for each participant. Data in % of maximum score.*

**Non-Useful Tetriminos Received events: Tests of Normality**

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
NonUseful_Arousal	.065	44	.200*	.980	44	.646
NonUseful_Valence	.076	44	.200*	.978	44	.556
NonUseful_Fun	.109	44	.200*	.981	44	.672
NonUseful_Excitement	.056	44	.200*	.988	44	.933
NonUseful_Boredom	.090	44	.200*	.970	44	.293
NonUseful_Frustration	.064	44	.200*	.986	44	.855
NonUseful_Challenge	.081	44	.200*	.987	44	.903

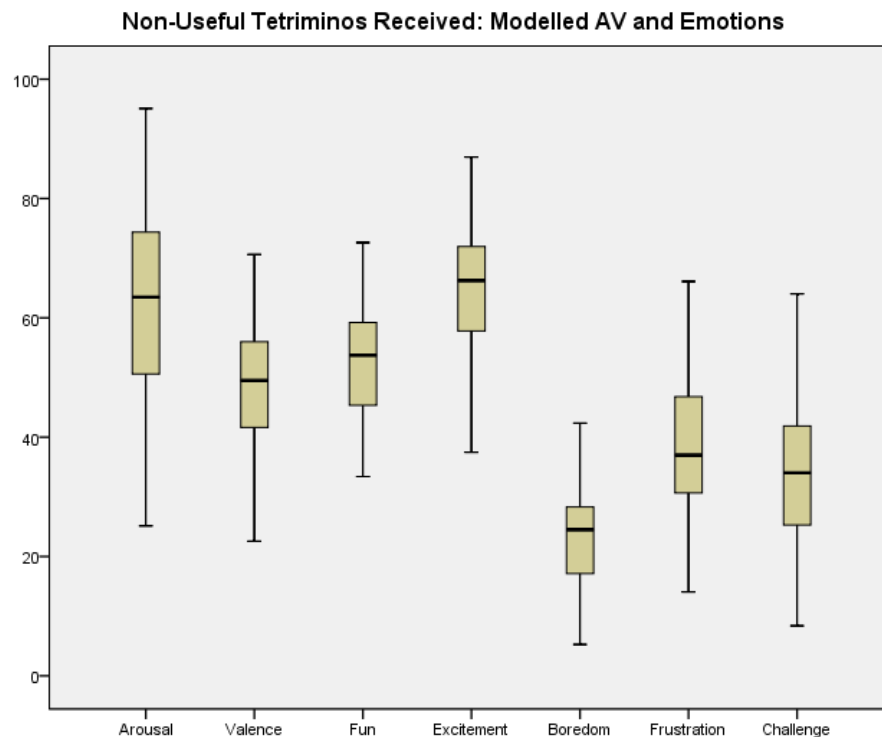
\*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

*Table 26: Results of Shapiro-Wilk test for normality, conducted for modelled AV space and emotions. All data was found to be normally distributed.*

As illustrated in Figure 52 below, no outliers were evident upon inspection of a box plot for values greater than 1.5 box lengths from the edge of the box.





*Figure 52: Box plot illustrating the distribution of AV space and emotions occurring during Non-Useful Tetriminos Received events. Note that no outliers are identified.*

The modelled arousal and valence scores for Non-Useful Tetriminos Received events were mapped onto the Affect Grid, illustrated in Figure 53 below, demonstrating that participant arousal was low-level and valence was low-negative. The standard deviation for both arousal and valence is represented through the proportionally skewed outer ellipse.

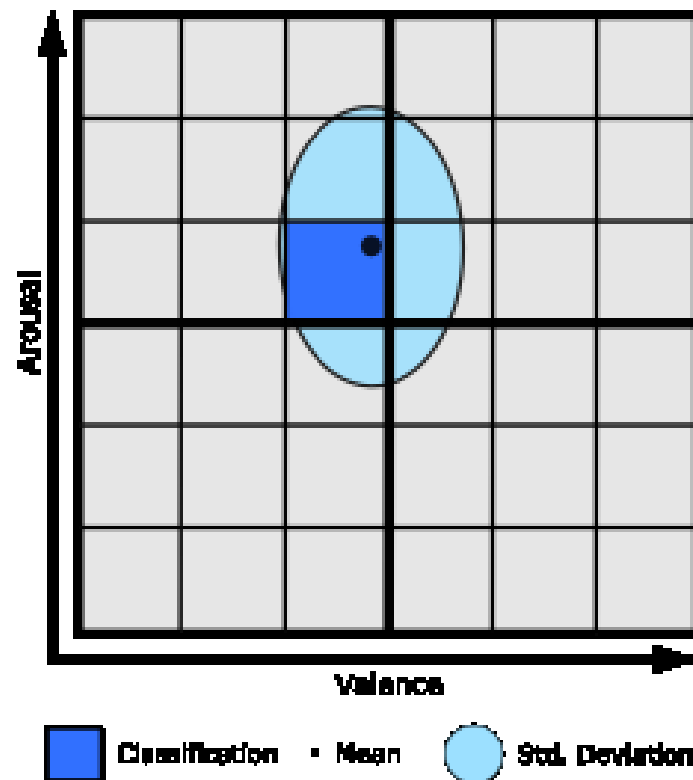


Figure 53: Affect Grid modelling the mean arousal and valence scores occurring during Non-Useful Tetriminos Received events, illustrating that the affective responses were minimal but slightly negatively valenced.

Paired-sample t-tests comparing means for modelled emotions against a baseline of the overall game play experience revealed that Non-Useful Tetriminos Received events elicited a range of influences upon modelled emotions, as illustrated in Table 27 below:

- **Decreased Fun:** A medium strength (*Cohen's*  $d = .55$ ) statistically **significant** mean decrease of -7.99 (95% CI, -12.41 to -3.58) in modelled fun compared to the overall game experience,  $t(43) = -3.65, p = 0.001$ .
- **Similar Excitement:** A small strength (*Cohen's*  $d = .23$ ) statistically **insignificant** mean increase of 2.12 (95% CI, 1.12 to 8.43) in modelled excitement compared to the overall game experience,  $t(43) = 3.95, p < .153$ .
- **Similar Boredom:** A trivial strength (*Cohen's*  $d = .03$ ) statistically **insignificant** mean increase of .37 (95% CI, -3.42 to 4.17) in modelled boredom compared to the overall game experience,  $t(43) = .196, p = .846$ .

- **Increased Frustration:** A large strength (*Cohen's d* = 1.08) statistically **significant** mean increase of 16.97 (95% CI, 12.21 to 21.74) in modelled frustration compared to the overall game experience,  $t(43) = 7.183$ ,  $p < .0005$ .
- **Similar Challenge:** A trivial strength (*Cohen's d* = .04) statistically **insignificant** mean decrease of -.66 (95% CI, -5.97 to 4.65) in modelled challenge compared to the overall game experience,  $t(43) = -2.50$ ,  $p = .81$ .

### Paired Samples Test: Non-Useful Tetriminos Received Events vs. Mean Game Play

		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	NonUseful_Fun - Combined_Fun	-7.99	14.52	2.19	-12.41	-3.58	-3.65	43	.001
Pair 2	NonUseful_Excitement - Combined_Excitement	2.12	17.15	2.59	1.12	8.43	3.95	43	.153
Pair 3	NonUseful_Boredom - Combined_Boredom	.37	12.49	1.88	-3.43	4.17	.196	43	.846
Pair 4	NonUseful_Frustration - Combined_Frustration	16.97	15.67	2.36	12.21	21.74	7.182	43	< .0005
Pair 5	Tetris_Challenge - Combined_Challenge	-.66	17.46	2.63	-5.97	4.65	-2.503	43	.81

Table 27: Results of paired-sample t-tests comparing mean Non-Useful Tetriminos Received events against mean overall game experience. Mean differences and the significance of these are demonstrated for each of the five modelled emotions.

#### 4.2.2.2.2. Correct Tetriminos Placement

During correct Tetriminos placement game events, the mean arousal and valence were 72.85% (+ 11.09%) and 63.22% (+ 12.09%) respectively. The results for each of the five modelled emotions are illustrated in Table 28 below. A Shapiro-Wilk test

for normality was conducted for all data, with all passing and therefore being normally distributed (see Table 29 below).

**Correct Tetriminos Placement Events: Descriptive Statistics**

	N	Mean	Std. Deviation
CorrectPlace_Arousal	44	72.85	11.09
CorrectPlace_Valence	44	63.22	12.10
CorrectPlace_Fun	44	69.56	13.12
CorrectPlace_Excitement	44	66.49	14.46
CorrectPlace_Boredom	44	14.28	5.03
CorrectPlace_Frustration	44	16.67	7.70
CorrectPlace_Challenge	44	27.17	10.29

Table 28: Descriptive statistics for modelled AV space and emotions, averaged across each time Tetriminos were placed correctly for each participant. Data in % of maximum score.

**Correct Tetriminos Placement events: Tests of Normality**

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
CorrectPlace_Arousal	.077	44	.200*	.981	44	.687
CorrectPlace_Valence	.141	44	.028	.966	44	.225
CorrectPlace_Fun	.077	44	.200*	.969	44	.284
CorrectPlace_Excitement	.082	44	.200*	.974	44	.422
CorrectPlace_Boredom	.105	44	.200*	.970	44	.305
CorrectPlace_Frustration	.125	44	.082	.945	44	.135
CorrectPlace_Challenge	.096	44	.200*	.968	44	.247

Table 29: Results of Shapiro-Wilk test for normality, conducted for modelled AV space and emotions. All data was found to be normally distributed.

As illustrated in Figure 54 below, no outliers were evident upon inspection of a box plot for values greater than 1.5 box lengths from the edge of the box.

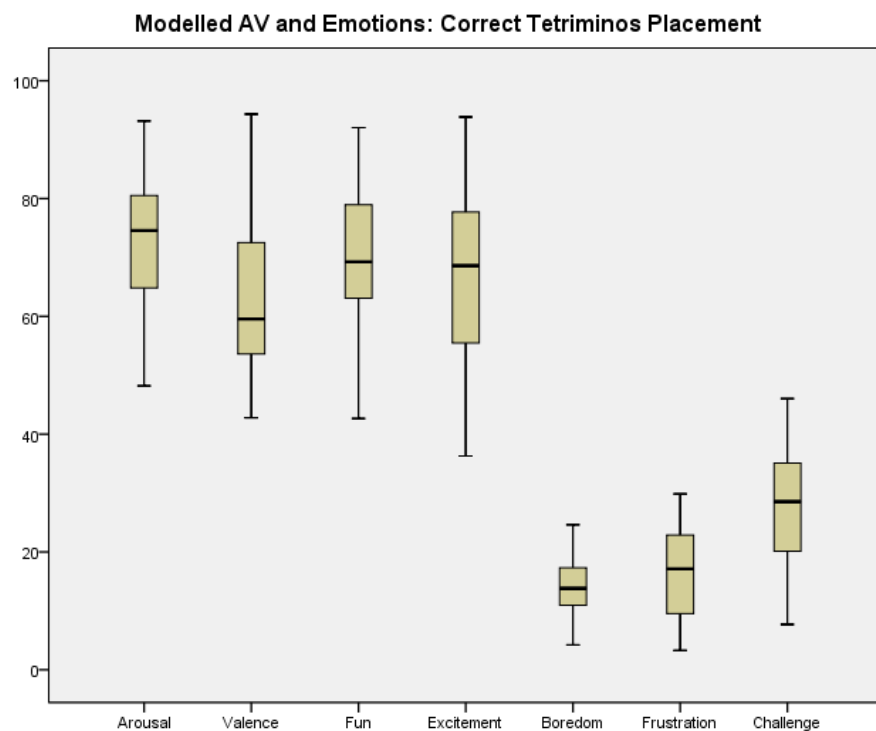


Figure 54: Box plot illustrating the distribution of AV space and emotions occurring during Correct Tetriminos Placement events. Note that no outliers are identified.

The modelled arousal and valence scores for Correct Tetriminos Placement events were mapped onto the Affect Grid, illustrated in Figure 55 below, demonstrating that participant arousal was mid-level and valence was low-positive. The standard deviation for both arousal and valence is represented through the proportionally skewed outer ellipse.

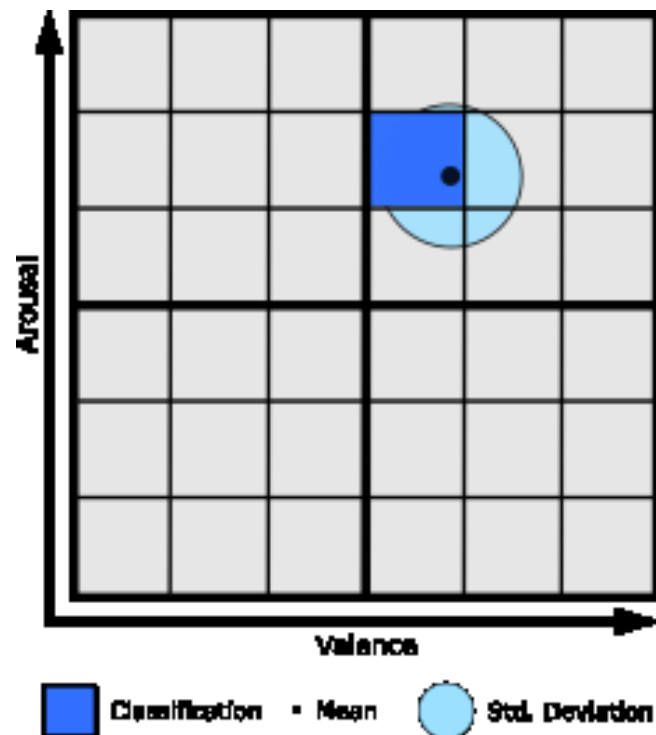


Figure 55: Affect Grid modelling the mean arousal and valence scores occurring during Correct Tetriminos Placement events, illustrating that the affective responses were slightly moderately arousing and positively valenced.

Paired-sample t-tests comparing means for modelled emotions against a baseline of the overall game play experience revealed that Correct Tetriminos Placement events elicited a range of influences upon modelled emotions, as illustrated in Table 30 below:

- **Increased Fun:** A medium strength (*Cohen's d* = .56) statistically **significant** mean increase of 8.49% (95% CI, 3.86 to 13.13) in modelled fun compared to the overall game experience,  $t(43) = 3.69, p = .001$ .
- **Increased Excitement:** A medium strength (*Cohen's d* = .52) statistically **significant** mean increase of 11.44 (95% CI, 4.74 to 18.14) in modelled excitement compared to the overall game experience,  $t(43) = 3.44, p = .001$ .
- **Decreased Boredom:** A large strength (*Cohen's d* = .84) statistically **significant** mean decrease of -8.77 (95% CI, -11.94 to -5.60) in modelled boredom compared to the overall game experience,  $t(43) = -5.578, p < .0005$ .

- **Decreased Frustration:** A medium strength (*Cohen's d* = .42) statistically **significant** mean decrease of -4.54 (95% CI, -7.85 to -1.24) in modelled frustration compared to the overall game experience,  $t(43) = -2.775$ ,  $p = .008$ .
- **Decreased Challenge:** A medium strength (*Cohen's d* = .49) statistically **significant** mean decrease of -7.39 (95% CI, -11.95 to -2.84) in modelled challenge compared to the overall game experience,  $t(43) = -3.28$ ,  $p = .002$ .

### Paired Samples Test: Correct Tetriminos Placement Events vs. Mean Game Play

		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2- tailed)
					Lower	Upper			
Pair 1	CorrectPlace_Fun - Combined_Fun	8.49	15.24	2.3	3.86	13.13	3.69	43	.001
Pair 2	CorrectPlace_Excitement - Combined_Excitement	11.44	22.04	3.32	4.74	18.14	3.44	43	.001
Pair 3	CorrectPlace_Boredom - Combined_Boredom	- 8.76607	10.42360	1.57142	- 11.93514	- 5.59701	- 5.578	43	< .0005
Pair 4	CorrectPlace_Frustration - Combined_Frustration	- 4.54382	10.86008	1.63722	-7.84559	- 1.24206	- 2.775	43	.008
Pair 5	CorrectPlace_Challenge - Combined_Challenge	-7.39	14.97	2.56	-11.95	-2.84	-3.28	43	.002

Table 30: Results of paired-sample t-tests comparing mean Correct Tetriminos Placement events against mean overall game experience. Mean differences and the significance of these are demonstrated for each of the five modelled emotions.

#### 4.2.2.2.3. Incorrect Tetriminos Placement

During Incorrect Tetriminos Placement events, the mean arousal and valence were 71.24% (+/- 14.48%) and 40.94% (+/- 15.21%) respectively. The results for each of the five modelled emotions are illustrated in Table 31 below. A Shapiro-Wilk test for normality was conducted for all data, with all passing and therefore being normally distributed (see Table 32 below).

**Incorrect Tetriminos Placement Events: Descriptive Statistics**

	N	Mean	Std. Deviation
IncorrectPlace_Arousal	44	71.24	14.48
IncorrectPlace_Valence	44	40.94	15.21
IncorrectPlace_Fun	44	56.90	15.81
IncorrectPlace_Excitement	44	66.83	15.52
IncorrectPlace_Boredom	44	18.00	6.01
IncorrectPlace_Frustration	44	48.33	11.02
IncorrectPlace_Challenge	44	52.38	12.38

Table 31: Descriptive statistics for modelled AV space and emotions, averaged across each time Tetriminos were incorrectly placed for each participant. Data in % of maximum score.

**Incorrect Tetriminos Placement events: Tests of Normality**

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
IncorrectPlace_Arousal	.069	44	.200*	.977	44	.503
IncorrectPlace_Valence	.070	44	.200*	.991	44	.977
IncorrectPlace_Fun	.090	44	.200*	.967	44	.245
IncorrectPlace_Excitement	.111	44	.200*	.950	44	.154
IncorrectPlace_Boredom	.112	44	.200*	.961	44	.142
IncorrectPlace_Frustration	.121	44	.111	.972	44	.369
IncorrectPlace_Challenge	.085	44	.200*	.969	44	.288

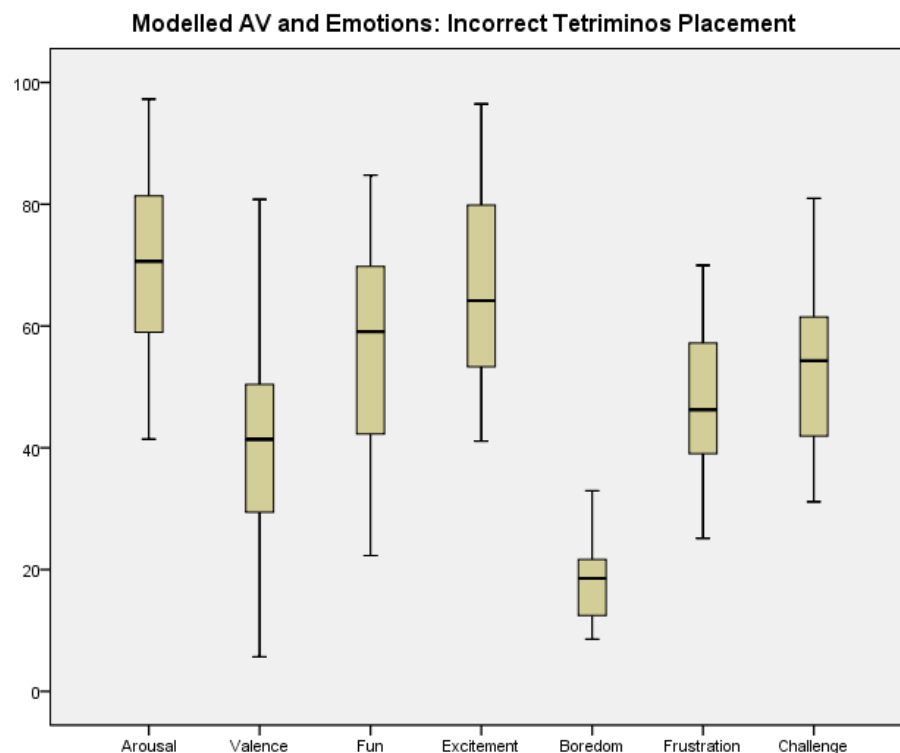
\*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Table 32: Results of Shapiro-Wilk test for normality, conducted for modelled AV space and emotions. All data was found to be normally distributed.

As illustrated in Figure 56 below, no outliers were evident upon inspection of a box plot for values greater than 1.5 box lengths from the edge of the box.





*Figure 56: Box plot illustrating the distribution of AV space and emotions occurring during Incorrect Tetriminos Placement events. Note that no outliers were identified.*

The modelled arousal and valence scores for Incorrect Tetriminos Placement events were mapped onto the Affect Grid, illustrated in Figure 57 below, demonstrating that participant arousal was mid-level and valence was low-negative. The standard deviation for both arousal and valence is represented through the proportionally skewed outer ellipse.

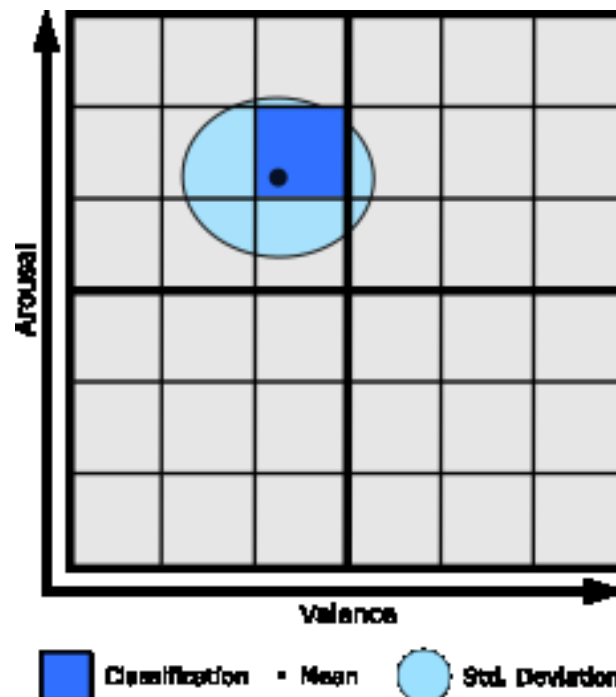


Figure 57: Affect Grid modelling the mean arousal and valence scores occurring during Incorrect Tetriminos Placement events, illustrating that the affective responses were minimal but slightly negatively valenced.

Paired-sample t-tests comparing means for modelled emotions against a baseline of the overall game play experience revealed that Incorrect Tetriminos Placement events elicited a range of influences upon modelled emotions, as illustrated in Table 33 below:

- **Similar Fun:** A small strength (*Cohen's d* = .20) statistically *insignificant* mean decrease of -4.18 (95% CI, -10.57 to 2.21) in modelled fun compared to the overall game experience,  $t(43) = -1.32, p = .195$ .
- **Increased Excitement:** A medium strength (*Cohen's d* = .59) statistically *significant* mean increase of 11.78 (95% CI, 5.68 to 17.87) in modelled excitement compared to the overall game experience,  $t(43) = 3.90, p < .0005$ .
- **Decreased Boredom:** A medium strength (*Cohen's d* = .51) statistically *significant* mean decrease of -5.04 (95 CI, -8.03% to -2.05) in modelled boredom compared to the overall game experience,  $t(43) = -3.402, p = .001$ .

- Increased Frustration:** A large strength (*Cohen's d* = 1.99) statistically **significant** mean increase of 27.12 (95% CI, 22.99 to 31.25) in modelled frustration compared to the overall game experience,  $t(43) = 13.261$ ,  $p < .0005$ .
- Increased Challenge:** A large strength (*Cohen's d* = 1.13) statistically **significant** mean increase of 17.81 (95% CI, 13.01 to 22.61) in modelled challenge compared to the overall game experience,  $t(43) = 7.49$ ,  $p < .0005$ .

### Paired Samples Test: Incorrect Tetriminos Placement Events vs. Mean Game Play

		Paired Differences							
			Std.	Std.	95% Confidence Interval of the Difference			Sig. (2-tailed)	
		Mean	Deviation	Error Mean	Lower	Upper	t	df	
Pair 1	IncorrectPlace_Fun - Combined_Fun	-4.18	21.12	3.17	-10.57	2.21	-1.32	43	.195
Pair 2	IncorrectPlace_Excitement - Combined_Excitement	11.78	20.05	3.02	5.68	17.87	3.90	43	< .0005
Pair 3	IncorrectPlace_Boredom - Combined_Boredom	-5.04	9.831	1.48	-8.031	-2.05	-3.402	43	.001
Pair 4	IncorrectPlace_Frustration - Combined_Frustration	27.12	13.57	2.04	22.99	31.24	13.261	43	< .0005
Pair 5	IncorrectPlace_Challenge - Combined_Challenge	17.81	15.78	2.38	13.01	22.61	7.49	43	< .0005

Table 34: Results of paired-sample t-tests comparing mean Incorrect Tetriminos Placement events against mean overall game experience. Mean differences and the significance of these are demonstrated for each of the five modelled emotions.

#### 4.2.2.3. Event Category: Player Performance

Player Performance includes events that relate to the performance or achievement of the player (coded as “Multiple Line Clear (“Tetris”)”, “Player Win”, and “Player Loss”). In the case of “Multiple Line Clear (“Tetris”)”, this can occur multiple times during game play depending upon the relative skill of the player. In the case of “Player Win” and “Player Loss”, these occur at the conclusion of each game play

session, with either one or the other occurring depending upon whether the player won or lost the game against their opponent.

#### 4.2.2.3.1. Multiple Line Clear ("Tetris")

During Multiple Line Clear ("Tetris") game events, the mean arousal and valence were 76.19% (+ 11.79%) and 69.92% (+ 13.17%) respectively. The results for each of the five modelled emotions are illustrated in table x below. A Shapiro-Wilk test for normality was conducted for all data, with all passing and therefore being normally distributed (see Table 35 below).

**Multiple Line Clear ("Tetris") Events: Descriptive Statistics**

	N	Mean	Std. Deviation
Tetris_Arousal	44	76.19	11.79
Tetris_Valence	44	69.92	13.17
Tetris_Fun	44	70.35	14.32
Tetris_Excitement	44	64.98	14.62
Tetris_Boredom	44	9.23	1.59
Tetris_Frustration	44	11.35	7.05
Tetris_Challenge	44	18.45	7.65

Table 35: Descriptive statistics for modelled AV space and emotions, averaged across each occurrence of a player clearing multiple lines for each participant. Data in % of maximum score.

**Multiple Line Clear ("Tetris") Events: Tests of Normality**

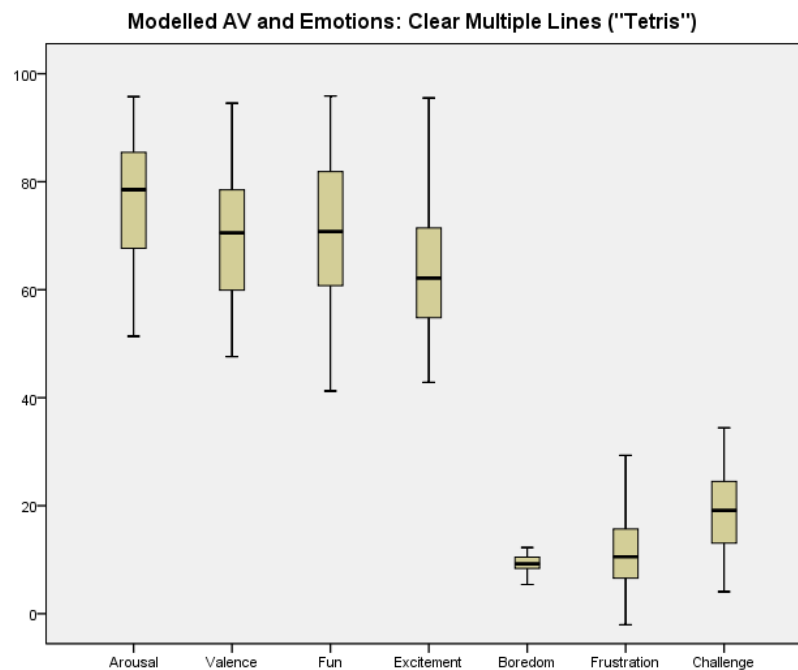
	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Tetris_Arousal	.098	44	.200*	.970	44	.311
Tetris_Valence	.076	44	.200*	.964	44	.189
Tetris_Fun	.073	44	.200*	.975	44	.443
Tetris_Excitement	.094	44	.200*	.969	44	.275
Tetris_Boredom	.091	44	.200*	.980	44	.640
Tetris_Frustration	.128	44	.068	.954	44	.177
Tetris_Challenge	.074	44	.200*	.981	44	.677

\*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Table 36: Results of Shapiro-Wilk test for normality, conducted for modelled AV space and emotions. All data was found to be normally distributed.

As illustrated in Figure 58 below, no outliers were evident upon inspection of a box plot for values greater than 1.5 box lengths from the edge of the box.



*Figure 58: Box plot illustrating the distribution of AV space and emotions occurring during Player Clears Multiple Lines ("Tetris") events. Note that no outliers are identified.*

The modelled arousal and valence scores for Multiple Line Clear ("Tetris") events were mapped onto the Affect Grid, illustrated in Figure 59 below, demonstrating that participant arousal was mid-level and valence was mid-positive. The standard deviation for both arousal and valence is represented through the proportionally skewed outer ellipse.

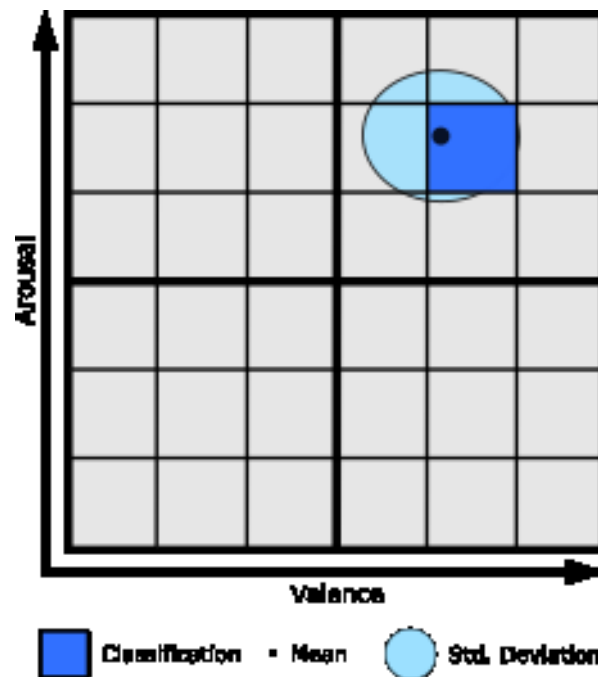


Figure 59: Affect Grid modelling the mean arousal and valence scores occurring during Multiple Line Clear ("Tetris") events, illustrating that the affective responses were minimal but slightly negatively valenced.

Paired-sample t-tests comparing means for modelled emotions against a baseline of the overall game play experience revealed that Multiple Line Clear ("Tetris") events elicited a range of influences upon modelled emotions, as illustrated in Table 37 below:

- **Increased Fun:** A medium strength (*Cohen's d* = .48) statistically **significant** mean increase of 9.27 (95% CI, 3.34 to 15.20) in modelled fun compared to the overall game experience,  $t(43) = 3.15, p = .003$ .
- **Increased Excitement:** A medium strength (*Cohen's d* = .52) statistically **significant** mean increase of 9.93 (95% CI, 4.07 to 15.78) in modelled excitement compared to the overall game experience,  $t(43) = 3.42, p < .0005$ .
- **Decreased Boredom:** A large strength (*Cohen's d* = 1.55) statistically **significant** mean decrease of -13.81 (95% CI, -16.51 to -11.11) in modelled boredom compared to the overall game experience,  $t(43) = -10.311, p < .0005$ .

- **Decreased Frustration:** A large strength (*Cohen's d* = .93) statistically **significant** mean decrease of -9.86 (95% CI, -13.07 to -6.65) in modelled frustration compared to the overall game experience,  $t(43) = -6.2, p < .0005$ .
- **Decreased Challenge:** A large strength (*Cohen's d* = 1.63) statistically **significant** mean decrease of -16.12 (95% CI, -19.12 to -13.11) in modelled challenge compared to the overall game experience,  $t(43) = -10.82, p < .0005$ .

#### Paired Samples Test: Multiple Line Clear ("Tetris") vs. Mean Game Play

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Tetris_Fun - Combined_Fun	9.27	19.51	2.94	3.34	15.20	3.15	43	.003
Pair 2	Tetris_Excitement - Combined_Excitement	9.93	19.25	2.90	4.07	15.78	3.42	43	< .0005
Pair 3	Tetris_Boredom - Combined_Boredom	-13.81	8.89	1.34	-16.51	-11.11	-10.31	43	< .0005
Pair 4	Tetris_Frustration - Combined_Frustration	-9.86	10.56	1.59	-13.07	-6.65	-6.20	43	< .0005
Pair 5	Tetris_Challenge - Combined_Challenge	-16.12	9.88	1.49	-19.12	-13.11	-10.82	43	< .0005

Table 37: Results of paired-sample t-tests comparing mean Multiple Line Clear ("Tetris") events against mean overall game experience. Mean differences and the significance of these are demonstrated for each of the five modelled emotions.

#### 4.2.2.3.2. Player Win

During Player Win events, the mean arousal and valence were 82.03% (+/- 10.12%) and 68.53% (+/- 10.76%) respectively. The results for each of the five modelled emotions are illustrated in Table 38 below. A Shapiro-Wilk test for normality was conducted for all data, with all passing and therefore being normally distributed (see Table 39 below).

#### Player Win Events: Descriptive Statistics

	N	Mean	Std. Deviation
Win_Arousal	44	82.03	10.12

Win_Valence	43	68.53	10.76
Win_Fun	44	69.90	10.67
Win_Excitement	44	79.23	13.35
Win_Boredom	44	10.95	4.07
Win_Frustration	44	10.62	4.41
Win_Challenge	44	31.66	13.23

Table 38: Descriptive statistics for modelled AV space and emotions, averaged across Player Win events for each participant. Data in % of maximum score.

### Player Win Events: Tests of Normality

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Win_Arousal	.100	43	.200*	.961	43	.153
Win_Valence	.126	43	.084	.959	43	.124
Win_Fun	.090	43	.200*	.973	43	.410
Win_Excitement	.105	43	.200*	.941	43	.127
Win_Boredom	.100	43	.200*	.970	43	.313
Win_Frustration	.120	43	.134	.967	43	.243
Win_Challenge	.089	43	.200*	.968	43	.259

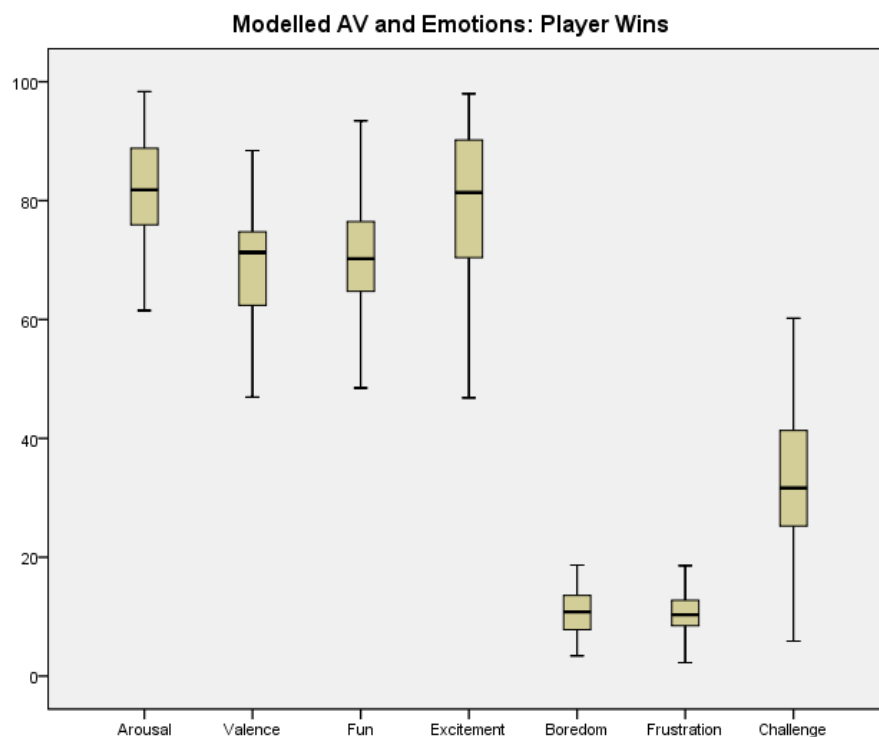
\*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Table 39: Results of Shapiro-Wilk test for normality, conducted for modelled AV space and emotions. All data was found to be normally distributed.

As illustrated in Figure 60 below, no outliers were evident upon inspection of a box plot for values greater than 1.5 box lengths from the edge of the box.





*Figure 60: Box plot illustrating the distribution of AV space and emotions occurring during Player Win events. Note that no outliers are identified.*

The modelled arousal and valence scores for Player Win events were mapped onto the Affect Grid, illustrated in Figure 61 below, demonstrating that participant arousal was high-level and valence was mid-positive. The standard deviation for both arousal and valence is represented through the proportionally skewed outer ellipse.

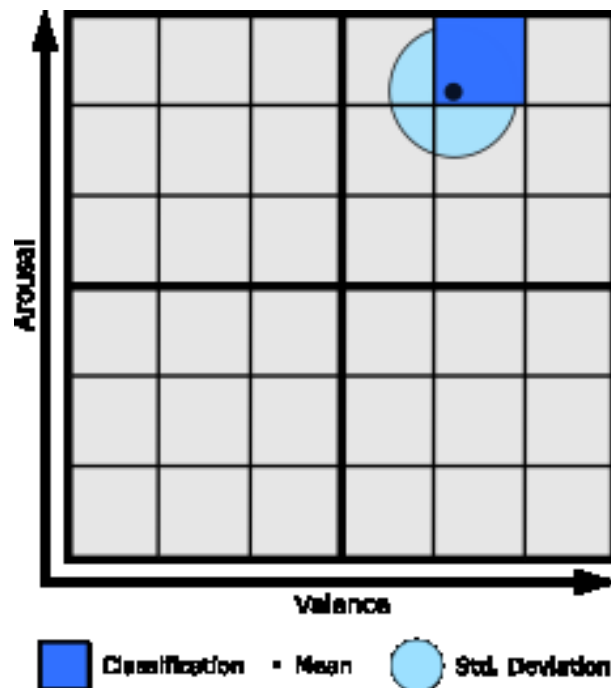


Figure 61: Affect Grid modelling the mean arousal and valence scores occurring during Player Win events, illustrating that the affective responses were highly arousing and positively valenced.

Paired-sample t-tests comparing means for modelled emotions against a baseline of the overall game play experience revealed that Player Win events elicited a range of influences upon modelled emotions, as illustrated in Table 40 below:

- **Increased Fun:** A medium strength (*Cohen's d* = .56) statistically **significant** mean increase of 8.83 (95% CI, 4.03 to 13.63) in modelled fun compared to the overall game experience,  $t(43) = 3.71, p = .001$ .
- **Increased Excitement:** A large strength (*Cohen's d* = 1.24) statistically **significant** mean increase of 24.17 (95% CI, 18.23 to 30.12) in modelled excitement compared to the overall game experience,  $t(43) = 8.20, p < .0005$ .
- **Decreased Boredom:** A large strength (*Cohen's d* = 1.35) statistically **significant** mean decrease of -12.09 (95% CI, -14.82 to -9.37) in modelled boredom compared to the overall game experience,  $t(43) = -8.954, p < .0005$ .
- **Decreased Frustration:** A large strength (*Cohen's d* = 1.14) statistically **significant** mean decrease of -10.59 (95% CI, -13.42 to -7.75) in modelled

frustration compared to the overall game experience,  $t(43) = -7.533$ ,  $p < .0005$ .

- **Decreased Challenge:** A small strength (*Cohen's d* = .17) statistically *insignificant* mean decrease of -2.91 (95% CI, -8.03 to 2.21) in modelled challenge compared to the overall game experience,  $t(43) = -1.15$ ,  $p = .258$ .

#### Paired Samples Test: Player Win Events vs. Mean Game Play

		Mean	Std. Deviation	Paired Differences		t	df	Sig. (2-tailed)
				Std. Error Mean	95% Confidence Interval of the Difference Lower Upper			
Pair 1	Win_Fun - Combined_Fun	8.83	15.78	2.38	4.03 13.63	3.71	43	.001
Pair 2	Win_Excitement - Combined_Excitement	24.17	19.56	2.95	18.23 30.12	8.20	43	< .0005
Pair 3	Win_Boredom - Combined_Boredom	-12.09	8.96	1.35	-14.82 -9.37	-8.95	43	< .0005
Pair 4	Win_Frustration - Combined_Frustration	-10.59	9.32	1.41	-13.42 -7.75	-7.53	43	< .0005
Pair 5	Win_Challenge - Combined_Challenge	-2.91	16.86	2.54	-8.03 2.21	-1.15	43	.258

Table 40: Results of paired-sample t-tests comparing mean Player Win events against mean overall game experience. Mean differences and the significance of these are demonstrated for each of the five modelled emotions.

#### 4.2.2.3.3. Player Loss

During Player Loss game events, the mean arousal and valence were 70.89% (+/- 13.74%) and 53.49% (+/- 20.64%) respectively. The results for each of the five modelled emotions are illustrated in Table 41 below. A Shapiro-Wilk test for normality was conducted for all data, with all passing and therefore being normally distributed (see Table 42 below).

#### Player Loss Events: Descriptive Statistics

	N	Mean	Std. Deviation
Loss_Arousal	44	70.89	13.74

Loss_Valence	44	53.49	20.64
Loss_Fun	44	42.76	15.97
Loss_Excitement	44	59.54	19.37
Loss_Boredom	44	18.70	6.52
Loss_Frustration	44	38.29	10.61
Loss_Challenge	44	28.34	12.91

Table 41: Descriptive statistics for modelled AV space and emotions, averaged across Player Loss events for each participant. Data in % of maximum score.

### Player Loss Events: Tests of Normality

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Loss_Arousal	.094	44	.200*	.967	44	.228
Loss_Valence	.071	44	.200*	.986	44	.877
Loss_Fun	.096	44	.200*	.980	44	.620
Loss_Excitement	.100	44	.200*	.975	44	.461
Loss_Boredom	.140	44	.029	.972	44	.357
Loss_Frustration	.066	44	.200*	.980	44	.618
Loss_Challenge	.068	44	.200*	.982	44	.730

\*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Table 42: Results of Shapiro-Wilk test for normality, conducted for modelled AV space and emotions. All data was found to be normally distributed.

As illustrated in Figure 62 below, no outliers were evident upon inspection of a box plot for values greater than 1.5 box lengths from the edge of the box.

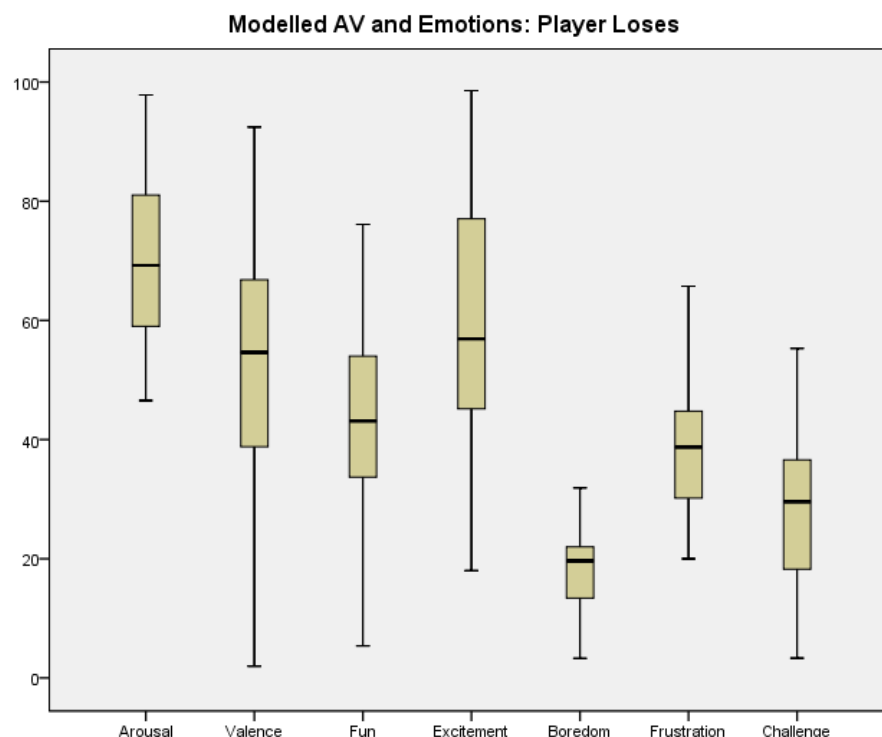


Figure 62: Box plot illustrating the distribution of AV space and emotions occurring during Player Loss events. Note that no outliers are identified.

The modelled arousal and valence scores for Player Loss events were mapped onto the Affect Grid, illustrated in Figure 63 below, demonstrating that participant arousal was mid-level and valence was very low-positive. The standard deviation for both arousal and valence is represented through the proportionally skewed outer ellipse.

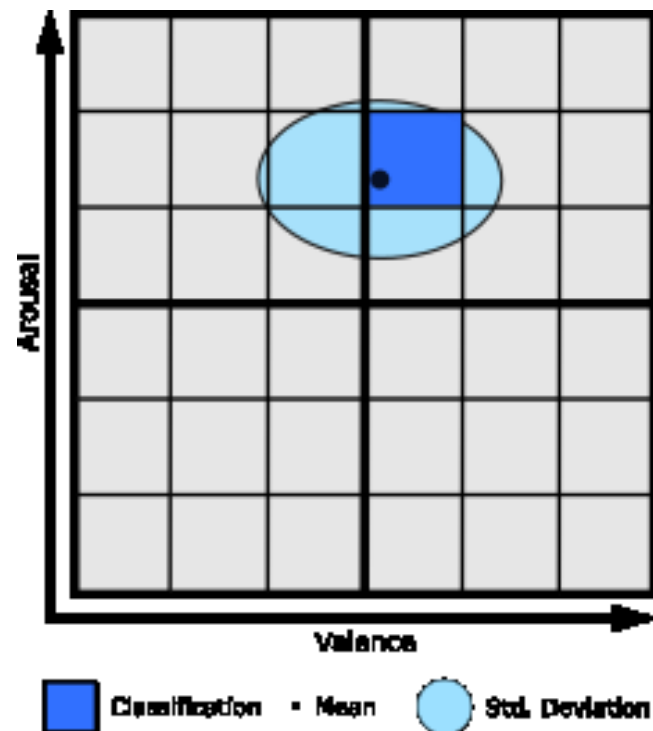


Figure 63: Affect Grid modelling the mean arousal and valence scores occurring during Player Loss events, illustrating that the affective responses were arousing and very low positively valenced.

Paired-sample t-tests comparing means for modelled emotions against a baseline of the overall game play experience revealed that Player Loss events elicited a range of influences upon modelled emotions, as illustrated in Table 43 below:

- **Decreased Fun:** A large strength (*Cohen's d* = .93) statistically **significant** mean decrease of -18.31 (95% CI, -24.34 to -12.29) in modelled fun compared to the overall game experience,  $t(43) = -6.13, p < .0005$ .
- **Similar Excitement:** A small strength (*Cohen's d* = .19) statistically **insignificant** mean increase of 4.49 (95% CI, -2.71 to 11.69) in modelled excitement compared to the overall game experience,  $t(43) = 1.26, p = .215$ .
- **Decreased Boredom:** A small strength (*Cohen's d* = .39) statistically **significant** mean decrease of -4.34 (95% CI, -7.76 to -.92) in modelled boredom compared to the overall game experience,  $t(43) = -2.560, p = .014$ .

- **Increased Frustration:** A large strength (*Cohen's d* = 1.18) statistically **significant** mean increase of 17.08 (95% CI, 12.66 to 21.50) in modelled frustration compared to the overall game experience,  $t(43) = 7.795$ ,  $p < .0005$ .
- **Decreased Challenge:** A medium strength (*Cohen's d* = .41) statistically **significant** mean decrease of -6.23 (95% CI, -10.96 to -1.50) in modelled challenge compared to the overall game experience,  $t(43) = -2.66$ ,  $p = .011$ .

**Paired Samples Test: Player Loss Events vs. Mean Game Play**

		Paired Differences							
			Std.	Std.	95% Confidence				
		Mean	Deviation	Error	Interval of the		t	df	Sig. (2-
				Mean	Lower	Upper			tailed)
Pair 1	Loss_Fun - Combined_Fun	-18.31	19.81	2.99	-24.34	-12.29	-6.13	43	< .0005
Pair 2	Loss_Excitement - Combined_Excitement	4.49	23.68	3.57	-2.71	11.69	1.26	43	.215
Pair 3	Loss_Boredom - Combined_Boredom	-4.34	11.25	1.70	-7.76	-.92	-2.56	43	.014
Pair 4	Loss_Frustration - Combined Frustration	17.08	14.54	2.19	12.66	21.50	7.79	43	< .0005
Pair 5	Loss_Challenge - Combined Challenge	-6.23	15.55	2.34	-10.96	-1.50	-2.66	43	.011

Table 43: Results of paired-sample t-tests comparing mean Player Loss events against mean overall game experience. Mean differences and the significance of these are demonstrated for each of the five modelled emotions.

#### 4.2.2.4. Event Category: Opponent Performance

Events that relate to player responses to the performance of their respective opponent (coded as "Opponent Correct Tetriminos Placement", "Opponent Incorrect Tetriminos Placement", and "Opponent Multiple Line Clear ("Tetris")") occur separately to the performance of the player themselves. This is because players cannot directly interact within Tetris Party Deluxe, however respective performances are visible to both players within the dyadic pair.

#### 4.2.2.4.1. Opponent Correct Tetriminos Placement

During Opponent Correct Tetriminos Placement game events, the mean arousal and valence were 62.32% (+/- 13.57%) and 56.56% (+/- 14.9%) respectively. The results for each of the five modelled emotions are illustrated in Table 44 below. A Shapiro-Wilk test for normality was conducted for all data, with all passing and therefore being normally distributed (see Table 45 below).

**Opponent Correct Tetriminos Placement Events: Descriptive Statistics**

	N	Mean	Std. Deviation
OppCorrectPlacement_Arousal	44	62.32	13.57
OppCorrectPlacement_Valence	44	56.56	14.90
OppCorrectPlacement_Fun	44	59.31	15.46
OppCorrectPlacement_Excitement	44	57.83	11.19
OppCorrectPlacement_Boredom	44	21.14	5.80
OppCorrectPlacement_Frustration	44	22.20	7.13
OppCorrectPlacement_Challenge	44	28.72	7.93

Table 44: Descriptive statistics for modelled AV space and emotions, averaged across occurrences of player opponents correctly placing Tetriminos for each participant. Data in % of maximum score.

**Opponent Correct Tetriminos Placement: Tests of Normality**

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
OppCorrectPlacement_Arousal	.082	44	.200*	.969	44	.291
OppCorrectPlacement_Valence	.049	44	.200*	.991	44	.976
OppCorrectPlacement_Fun	.072	44	.200*	.989	44	.939
OppCorrectPlacement_Excitement	.083	44	.200*	.977	44	.523
OppCorrectPlacement_Boredom	.066	44	.200*	.978	44	.569
OppCorrectPlacement_Frustration	.082	44	.200*	.984	44	.796
OppCorrectPlacement_Challenge	.067	44	.200*	.988	44	.920

\*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction



Table 45: Results of Shapiro-Wilk test for normality, conducted for modelled AV space and emotions. All data was found to be normally distributed.

As illustrated in Figure 64 below, no outliers were evident upon inspection of a box plot for values greater than 1.5 box lengths from the edge of the box.

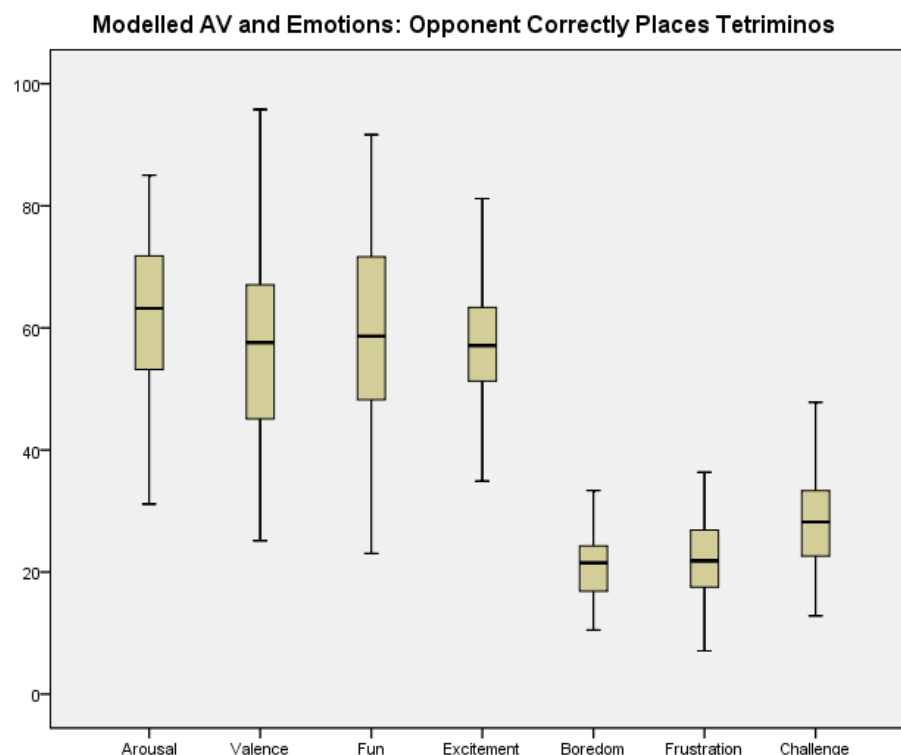


Figure 64: Box plot illustrating the distribution of AV space and emotions occurring during Opponent Correct Tetriminos Placement events. Note that no outliers are identified.

The modelled arousal and valence scores for Opponent Correct Tetriminos Placement events were mapped onto the Affect Grid, illustrated in Figure 65 below, demonstrating that participant arousal was low-level and valence was low-positive. The standard deviation for both arousal and valence is represented through the proportionally skewed outer ellipse.

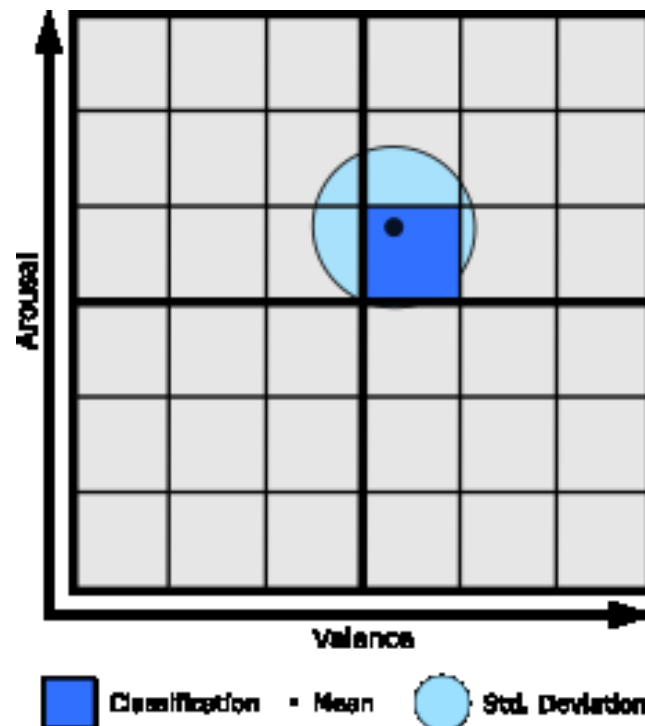


Figure 65: Affect Grid modelling the mean arousal and valence scores occurring during Opponent Correct Tetriminos Placement events, illustrating that the affective responses were minimal but slightly positively valenced.

Paired-sample t-tests comparing means for modelled emotions against a baseline of the overall game play experience revealed that Opponent Correct Tetriminos Placement events elicited a range of influences upon modelled emotions, as illustrated in Table 46 below:

- **Similar Fun:** A trivial strength (*Cohen's d* = .10) statistically **insignificant** mean decrease of -1.76 (95% CI, -7.20 to 3.69) in modelled fun compared to the overall game experience,  $t(43) = -.65$ ,  $p = .518$ .
- **Similar Excitement:** A small strength (*Cohen's d* = .16) statistically **insignificant** mean increase of 2.77 (95% CI, 2.50 to 8.05) in modelled excitement compared to the overall game experience,  $t(43) = 1.06$ ,  $p = .294$ .
- **Similar Boredom:** A small strength (*Cohen's d* = .17) statistically **insignificant** mean decrease of -1.90 (95% CI, -5.17 to -1.37) in modelled boredom compared to the overall game experience,  $t(43) = -1.173$ ,  $p = .247$ .

- **Similar Frustration:** A trivial strength (*Cohen's d* = .09) statistically *insignificant* mean increase of .99 (95% CI, -2.22 to 4.20) in modelled frustration compared to the overall game experience,  $t(43) = .622$ ,  $p = .537$ .
- **Decreased Challenge:** A medium strength (*Cohen's d* = .53) statistically *significant* mean decrease of -5.85 (95% CI, -9.23 to -2.46) in modelled challenge compared to the overall game experience,  $t(43) = -3.48$ ,  $p = .001$ .

#### Paired Samples Test: Player Response to Opponent Correctly Placing Tetriminos

		Paired Differences							
				Std.	95% Confidence Interval of the Difference				Sig. (2-tailed)
		Mean	Deviation	Error Mean	Lower	Upper	t	df	
Pair 1	OppCorrectPlacement_Fun - Combined_Fun	-1.76	17.90	2.70	-7.20	3.69	-.65	43	.518
Pair 2	OppCorrectPlacement_Excitement - Combined_Excitement	2.77	17.34	2.61	-2.50	8.05	1.06	43	.294
Pair 3	OppCorrectPlacement_Boredom - Combined_Boredom	-1.90	10.76	1.62	-5.17	1.37	-1.17	43	.247
Pair 4	OppCorrectPlacement_Frustration - Combined_Frustration	.99	10.56	1.59	-2.22	4.20	.62	43	.537
Pair 5	OppCorrectPlacement_Challenge - Combined_Challenge	-5.85	11.14	1.68	-9.23	-2.46	-3.48	43	.001

Table 46: Results of paired-sample t-tests comparing mean Opponent Correct Tetriminos Placement events against mean overall game experience. Mean differences and the significance of these are demonstrated for each of the five modelled emotions.

#### 4.2.2.4.2. Opponent Incorrect Tetriminos Placement

During Opponent Incorrect Tetriminos Placement game events, the mean arousal and valence were 68.89% (+ 13.57%) and 60.34% (+ 12.89%) respectively. The results for each of the five modelled emotions are illustrated in Table 47 below. A Shapiro-Wilk test for normality was conducted for all data, with all passing and therefore being normally distributed (see Table 48 below).

**Opponent Incorrect Tetriminos Placement Events: Descriptive Statistics**

	N	Mean	Std. Deviation
OpplIncorrectPlacement_Fun	44	67.04	17.42
OpplIncorrectPlacement_Excitement	44	51.33	10.20
OpplIncorrectPlacement_Boredom	44	14.34	4.45
OpplIncorrectPlacement_Frustration	44	9.08	2.32
OpplIncorrectPlacement_Challenge	44	16.10	4.91

Table 47: Descriptive statistics for modelled AV space and emotions, averaged per Opponent Incorrect Tetriminos Placement event. Data in % of maximum score.

**Opponent Incorrect Tetriminos Placement Events: Tests of Normality**

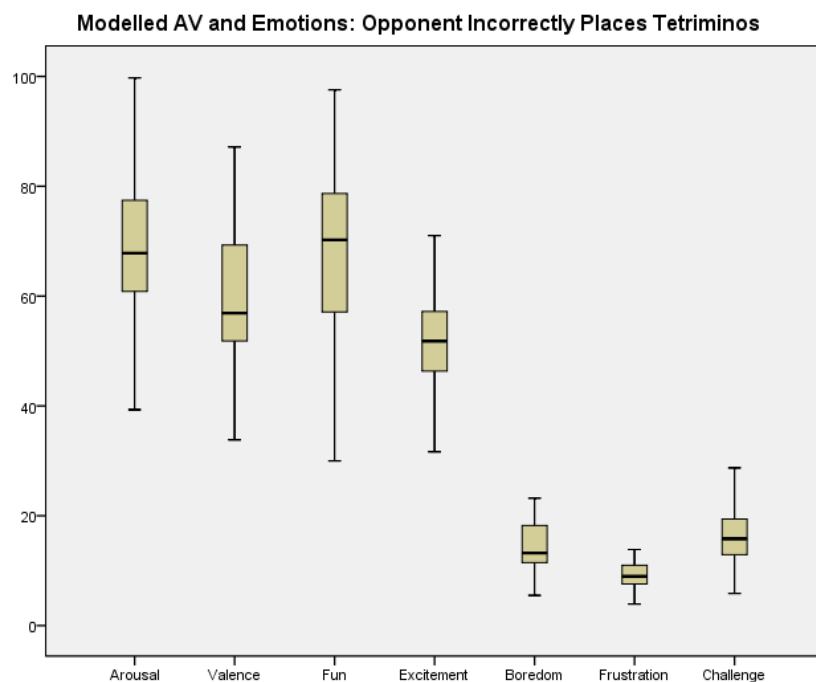
	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
OpplIncorrectPlacement_Arousal	.078	44	.200*	.981	44	.66
OpplIncorrectPlacement_Variance	.123	44	.091	.977	44	.57
OpplIncorrectPlacement_Fun	.097	44	.200*	.972	44	.34
OpplIncorrectPlacement_Excitement	.083	44	.200*	.970	44	.30
OpplIncorrectPlacement_Boredom	.137	44	.037	.969	44	.28
OpplIncorrectPlacement_Frustration	.083	44	.200*	.981	44	.66
OpplIncorrectPlacement_Challenge	.063	44	.200*	.992	44	.99

\*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Table 48 : Results of Shapiro-Wilk test for normality, conducted for modelled AV space and emotions. All data was found to be normally distributed.

As illustrated in Figure 66 below, no outliers were evident upon inspection of a box plot for values greater than 1.5 box lengths from the edge of the box.



*Figure 66: Box plot illustrating the distribution of AV space and emotions occurring during Opponent Incorrect Tetriminos Placement events. Note that no outliers are identified.*

The modelled arousal and valence scores for Opponent Incorrect Tetriminos Placement events were mapped onto the Affect Grid, illustrated in Figure 67 below, demonstrating that participant arousal was mid-level and valence was low-positive. The standard deviation for both arousal and valence is represented through the proportionally skewed outer ellipse.

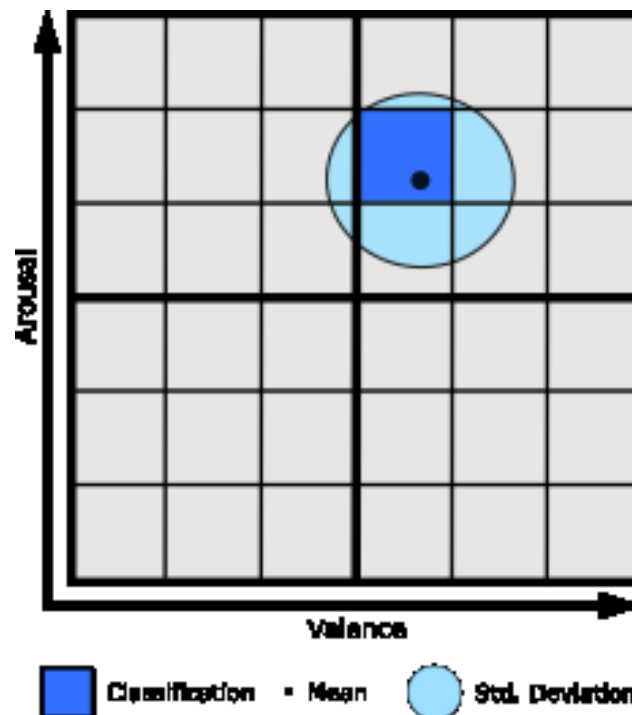


Figure 67: Affect Grid modelling the mean arousal and valence scores occurring during Opponent Incorrect Tetriminos Placement events, illustrating that the affective responses were moderate and slightly positively valenced.

Paired-sample t-tests comparing means for modelled emotions against a baseline of the overall game play experience revealed that Opponent Incorrect Tetriminos Placement events elicited a range of influences upon modelled emotions, as illustrated in Table 49 below:

- **Increased Fun:** A small strength (*Cohen's d* = .29) statistically **significant** mean increase of 5.97 (95% CI, -.26 to 12.20) in modelled fun compared to the overall game experience,  $t(43) = 1.93, p = .045$ .
- **Similar Excitement:** A small strength (*Cohen's d* = .22) statistically **insignificant** mean decrease of -3.72 (95% CI, -8.87 to 1.42) in modelled excitement compared to the overall game experience,  $t(43) = -1.46, p < .152$ .
- **Decreased Boredom:** A large strength (*Cohen's d* = .85) statistically **significant** mean decrease of -8.70 (95% CI, -11.81 to -5.59) in modelled boredom compared to the overall game experience,  $t(43) = -5.643, p < .0005$ .

- **Decreased Frustration:** A large strength (*Cohen's d* = 1.36) statistically **significant** mean decrease of -12.13 (95% CI, -14.84 to -9.41) in modelled frustration compared to the overall game experience,  $t(43) = -9.02, p < .0005$ .
- **Decreased Challenge:** A large strength (*Cohen's d* = 1.88) statistically **significant** mean decrease of -18.47 (95% CI, -21.48 to -15.48) in modelled challenge compared to the overall game experience,  $t(43) = -12.49, p < .0005$ .

#### Paired Samples Test: Player Response to Opponent Incorrectly Placing Tetriminos

		Paired Differences							
		Mean	Std. Deviation	Std. Error	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
			n	Mean	Lower	Upper			
Pair 1	OpplIncorrectPlacement_Fun - Combined_Fun	5.97	20.49	3.09	-.26	12.20	1.93	43	<b>.045</b>
Pair 2	OpplIncorrectPlacement_Excitement - Combined_Excitement	-3.72	16.93	2.55	-8.87	1.42	-1.46	43	.152
Pair 3	OpplIncorrectPlacement_Boredom - Combined_Boredom	-8.70	10.23	1.54	-11.81	-5.59	-5.64	43	<b>&lt; .0005</b>
Pair 4	OpplIncorrectPlacement_Frustration - Combined_Frustration	-12.13	8.92	1.34	-14.84	-9.41	-9.02	43	<b>&lt; .0005</b>
Pair 5	OpplIncorrectPlacement_Challenge - Combined_Challenge	-18.47	9.81	1.48	-21.45	-15.48	-12.49	43	<b>&lt; .0005</b>

Table 49: Results of paired-sample t-tests comparing mean Opponent Incorrect Tetriminos Placement events against mean overall game experience. Mean differences and the significance of these are demonstrated for each of the five modelled emotions.

#### 4.2.2.4.3. Opponent Multiple Line Clear ("Tetris")

During Opponent Multiple Line Clear ("Tetris") game events, the mean arousal and valence were 59.70% (+/- 18.18%) and 46.94% (+/- 15.98%) respectively. The results for each of the five modelled emotions are illustrated in Table 50 below. A Shapiro-

Wilk test for normality was conducted for all data, with all passing and therefore being normally distributed (see Table 51 below).

### Opponent Multiple Line Clear ("Tetris") Events: Descriptive Statistics

	N	Mean	Std. Deviation
OppTetris_Fun	44	60.5155	11.90600
OppTetris_Excitement	44	50.6945	11.76169
OppTetris_Boredom	44	13.6355	3.16292
OppTetris_Frustration	44	27.9141	11.29801
OppTetris_Challenge	44	34.0837	5.37271

Table 50: Descriptive statistics for modelled AV space and emotions, averaged across Opponent Multiple Line Clear ("Tetris") events for each participant. Data in % of maximum score.

### Opponent Multiple Line Clear ("Tetris") Events: Tests of Normality

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
OppTetris_Arousal	.111	44	.200*	.946	44	.138
OppTetris_Valence	.070	44	.200*	.987	44	.896
OppTetris_Fun	.101	44	.200*	.977	44	.519
OppTetris_Excitement	.091	44	.200*	.982	44	.713
OppTetris_Boredom	.077	44	.200*	.981	44	.686
OppTetris_Frustration	.120	44	.119	.977	44	.510
OppTetris_Challenge	.096	44	.200*	.984	44	.802

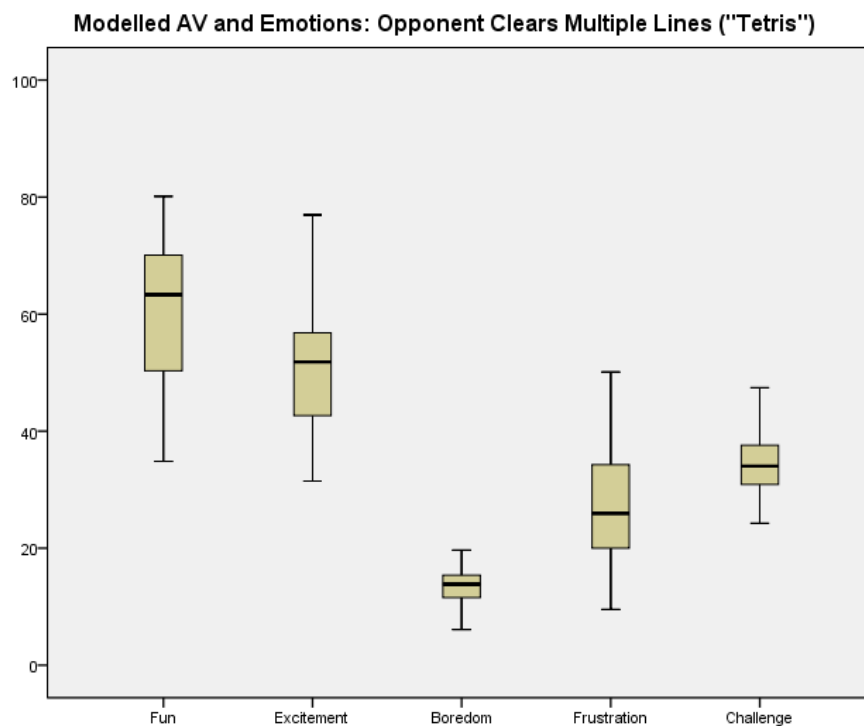
\*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Table 51: Results of Shapiro-Wilk test for normality, conducted for modelled AV space and emotions. All data was found to be normally distributed.

As illustrated in Figure 68 below, no outliers were evident upon inspection of a box plot for values greater than 1.5 box lengths from the edge of the box.





*Figure 68: Box plot illustrating the distribution of AV space and emotions occurring during Opponent Multiple Line Clear ("Tetris") events. Note that no outliers are identified.*

The modelled arousal and valence scores for Opponent Multiple Line Clear ("Tetris") events were mapped onto the Affect Grid, illustrated in Figure 69 below, demonstrating that participant arousal was low-level and valence was low-negative. The standard deviation for both arousal and valence is represented through the proportionally skewed outer ellipse.

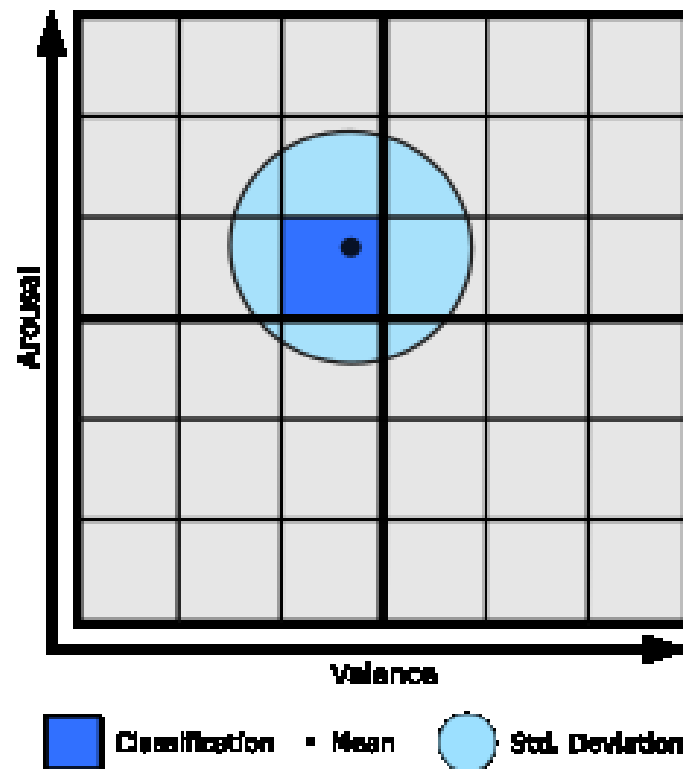


Figure 69: Affect Grid modelling the mean arousal and valence scores occurring during Opponent Multiple Line Clear (“Tetris”) events, illustrating that the affective responses were minimal but slightly negatively valenced.

Paired-sample t-tests comparing means for modelled emotions against a baseline of the overall game play experience revealed that Opponent Multiple Line Clear (“Tetris”) events elicited a range of influences upon modelled emotions, as illustrated in Table 52 below:

- **Similar Fun:** A trivial strength (*Cohen’s d* = 0.03) statistically **insignificant** mean decrease of -.56 (95% CI, -5.49 to 4.38) in modelled fun compared to the overall game experience,  $t(43) = -.23, p = .821$ .
- **Similar Excitement:** A small strength (*Cohen’s d* = .22) statistically **insignificant** mean decrease of -4.04 (95% CI, -9.57 to 1.49) in modelled excitement compared to the overall game experience,  $t(43) = -1.47, p = .148$ .
- **Decreased Boredom:** A large strength (*Cohen’s d* = 1.02) statistically **significant** mean decrease of -9.41 (95% CI, -12.21 to -6.61) in modelled boredom compared to the overall game experience,  $t(43) = -6.774, p < .0005$ .

- **Increased Frustration:** A medium strength (*Cohen's d* = .43) statistically **significant** mean increase of 6.36 (95% CI, 1.85 to 10.88) in modelled frustration compared to the overall game experience,  $t(43) = 2.84, p = .007$ .
- **Similar Challenge:** A trivial strength (*Cohen's d* = .04) statistically **insignificant** mean decrease of -.48 (95% CI, -3.86 to 2.89) in modelled challenge compared to the overall game experience,  $t(43) = -.29, p = .774$ .

**Paired Samples Test: Player Response when Opponent Clears Multiple Lines ("Tetris") vs. Mean Game Play**

		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	OppTetris_Fun - Combined_Fun	-.56	16.24	2.45	-5.49	4.38	-.23	43	.821
Pair 2	OppTetris_Excitement - Combined_Excitement	-4.04	18.19	2.74	-9.57	1.49	-1.47	43	.148
Pair 3	OppTetris_Boredom - Combined_Boredom	-9.41	9.21	1.39	-12.21	-6.61	-6.77	43	< .0005
Pair 4	OppTetris_Frustration - Combined_Frustration	6.36	14.84	2.24	1.85	10.88	2.84	43	.007
Pair 5	OppTetris_Challenge - Combined_Challenge	-.48	11.11	1.67	-3.86	2.89	-.29	43	.774

Table 52: Results of paired-sample t-tests comparing mean Opponent Multiple Line Clear ("Tetris") events against mean overall game experience. Mean differences and the significance of these are demonstrated for each of the five modelled emotions.

#### 4.2.3. Summary of Results for Research Question 1

The previous sections have presented the results of descriptive and inferential statistical analysis into the effects that events occurring during game play have upon the affective responses of players, and compare these to the mean game play experience. Research Question 1 sought to identify events occurring during video game play, and explore the influence that these can have upon affective responses.

As evidenced in Figure 70 and Figure 71 below, the ten game events included within this analysis (see Section 3.9 for details of event identification) elicited a range of influences upon the affective state of participants, the implications for which are discussed in Chapter 5: Discussions.

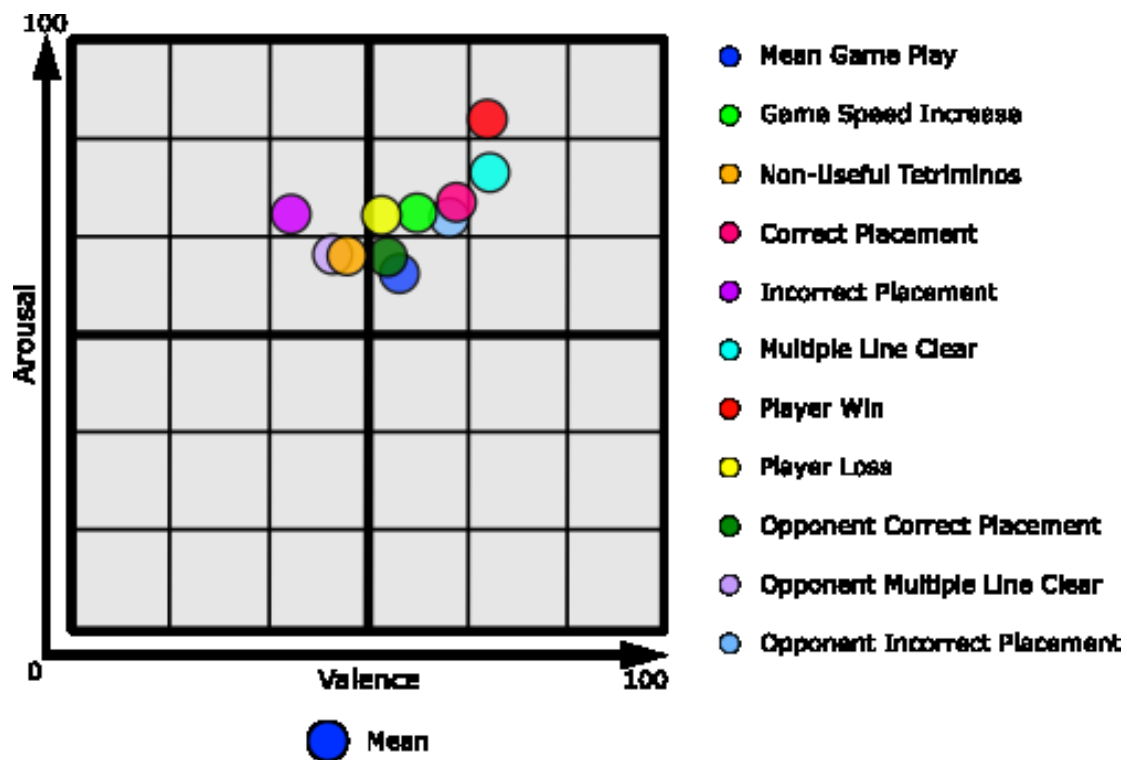
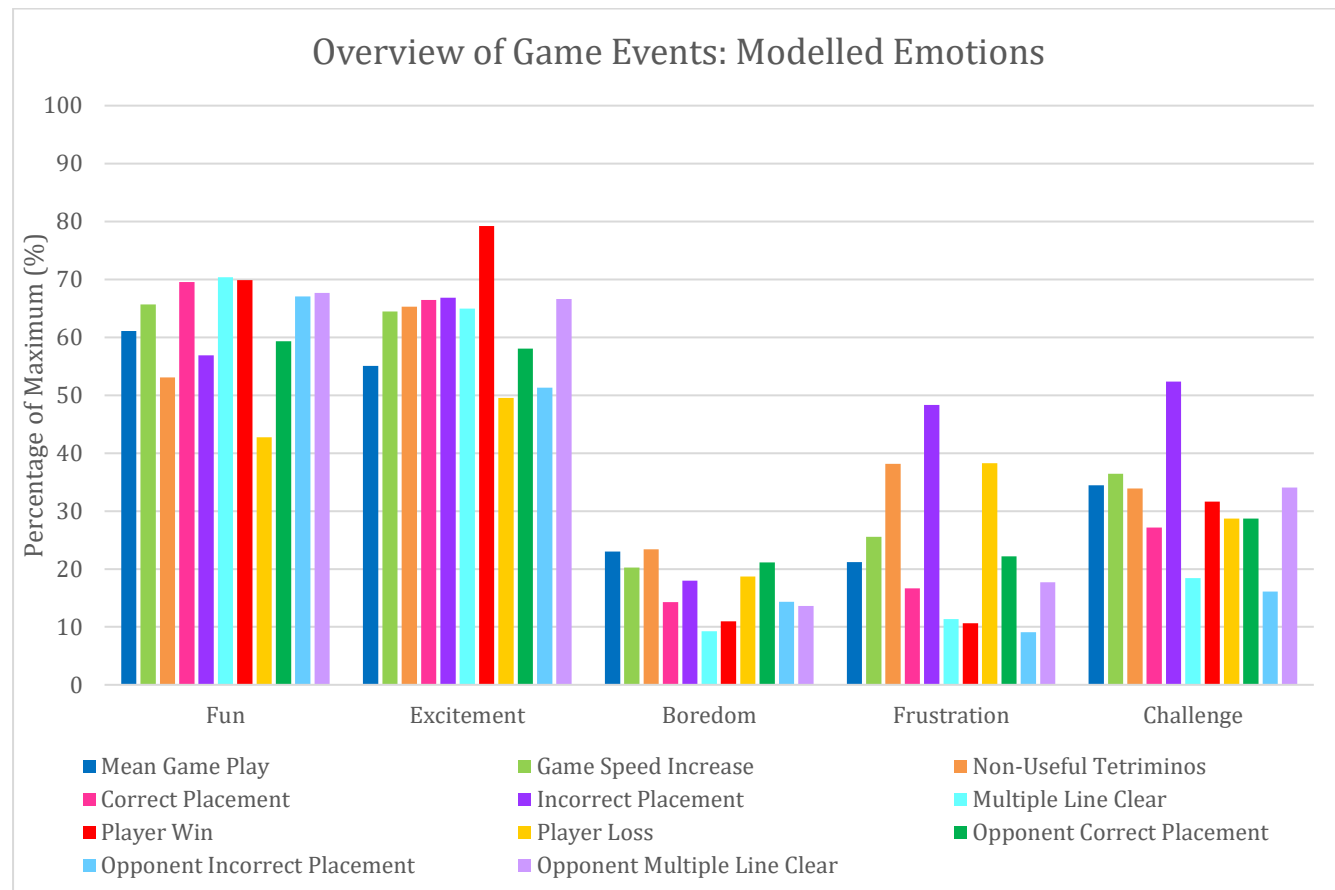


Figure 70: Overview of the modelled AV space for each of the game events analysed within the results section of this project. In-depth analysis, including effect size, strength, and significance of differences is presented in the previous sections.



*Figure 71: Overview of the modelled Emotional categories for each of the game events analysed within the results section of this project. In-depth analysis, including effect size, strength, and significance of differences is presented in the previous sections.*

### 4.3. Research Question 2

#### **Research Question 2:**

*How does the use of psychophysiological instruments influence the player experience during video game play?*

Research Question 2 aimed to explore the influence that the use of psychophysiological equipment, when placed upon the body of participants, had upon the game experience, both in terms of general perceived game experience, and affective experience. The unit of analysis was therefore player experience, analysing

the influence that physiological measurement devices had upon self-reported affect and game experience. This required the collection of subjective data pertaining to relevant game experience factors, namely the Game Experience Questionnaire (GEQ) to measure general perceived game experience, and the Self-Assessment Manikin (SAM) to measure affective experience.

Addressing Research Question 2 required manipulating the presence of psychophysiological equipment as the independent variable (IV) with two levels (Equipment On vs. Equipment Off), and the collection of subjective data relating to perceived game and emotional experiences as the dependent variables (DVs). The study manipulated equipment presence using a within-participant (repeated-measure) design (see Section 3.7.1 for details), and therefore the DVs are paired to participants, with each participant experiencing both Equipment On and Equipment Off conditions during their experimental session.

Considering the study design and data types used, the analysis of the influence of psychophysiological equipment therefore consists of:

- **Descriptive statistics:** Describing the survey results collected for overall mean survey responses, as well as describing the data when separated into Equipment On and Equipment Off experimental conditions. Additionally, data on observed facial expressions is included to provide further depth.
- **Paired-samples t-tests:** Paired-samples t-tests are commonly used to investigate whether mean differences between paired observations are statistically different from zero (Howell, 2012). Given that RQ2 requires an examination of whether the presence of psychophysiological equipment influences perceived game and affective experience, and with the study design and data types previously discussed, a paired-samples t-test is an appropriate test.

These address Research Question 2 through providing descriptive insights into the effects that psychophysiological equipment had upon subjective game experiences, as well as evidence to support the strength and significance of these effects.

### **4.3.1. Subjective Game Experience: Equipment On vs. Equipment Off**

Data on subjective game experience was collected through the Game Experience Questionnaire, which included a CORE component relating to the game play itself, and a SOCIAL component that related to social interactions within the game. CORE game experience results in percentages of maximum score for seven categories of game experience (Competence, Immersion, Flow, Tension, Challenge, Negative Affect, and Positive Affect), while SOCIAL game experience results in percentages of maximum score for three categories of social game experience (Empathy for opponent, Negative Feelings towards opponent, and Behavioural Involvement with opponent).

Descriptive statistics are provided for both CORE and SOCIAL components of subjective player experiences, measured for both Equipment On and Equipment Off conditions. Repeated-measures t-tests are then used to assess the effect size, strength, and significance of differences between the Equipment On and Equipment Off conditions for each aspect of subjective game experience.

A Shapiro-Wilk test for normality was conducted for all data, with all passing and therefore being normally distributed (see Table 53 below for example). Additionally, no outliers were evident in any of the data upon inspection of a box plot for values greater than 1.5 box-lengths from the edge of the box (see Figure 72 below for a sample).

**Sample Survey Data: Test of Normality**

	Statistic	Shapiro-Wilk df	Sig.
EQ Competence	.971	44	.333
EQ Immersion	.959	44	.117
EQ Flow	.968	44	.267
EQ Tension	.976	44	.473
EQ Challenge	.933	44	.135
EQ Negative Affect	.956	44	.193
EQ Positive Affect	.912	44	.126

Table 53: Results of Shapiro-Wilk test for normality, conducted in this case for the Game Experience Questionnaire CORE components under the Equipment On experimental condition. All data was found to be normally distributed, including those for data not presented.

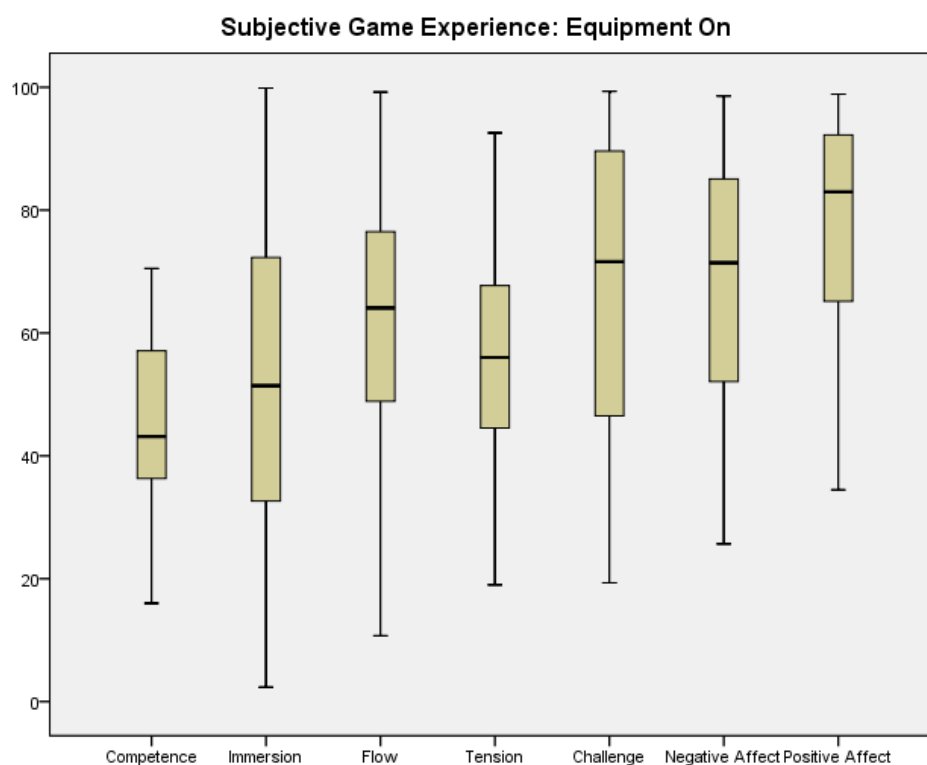


Figure 72: Box plot illustrating the distribution of subjective Game Experience Questionnaire CORE components. Note that no outliers are identified, including those for data not presented.



#### 4.3.1.1. Subjective CORE Game Experience Comparisons

The results for each of the seven components of the Game Experience Questionnaire CORE (GEQ: CORE) were averaged across all participants, separated into the Equipment On and Equipment Off experimental conditions. Survey data is presented as percentage of maximum score for each component. The mean data for both Equipment On and Equipment Off conditions are presented in Table 54 and Figure 73 below.

#### Game Experience Questionnaire CORE (GEQ: CORE) Survey Results:

		Equip. On	Equip. Off
<b>Competence</b>	Mean	45.33	51.92
	Std. Deviation	13.33	23.33
<b>Immersion</b>	Mean	54.17	69.83
	Std. Deviation	24.77	22.47
<b>Flow</b>	Mean	62.70	58.02
	Std. Deviation	22.72	24.19
<b>Tension</b>	Mean	55.71	56.96
	Std. Deviation	18.84	21.83
<b>Challenge</b>	Mean	67.16	58.05
	Std. Deviation	23.12	22.84
<b>Negative Affect</b>	Mean	68.51	54.88
	Std. Deviation	19.86	22.33
<b>Positive Affect</b>	Mean	77.79	58.03
	Std. Deviation	17.05	17.36

Table 54: Descriptive statistics for Game Experience Questionnaire CORE components, averaged across all participants, and separated into Equipment On and Equipment Off conditions. Data in % of maximum score.

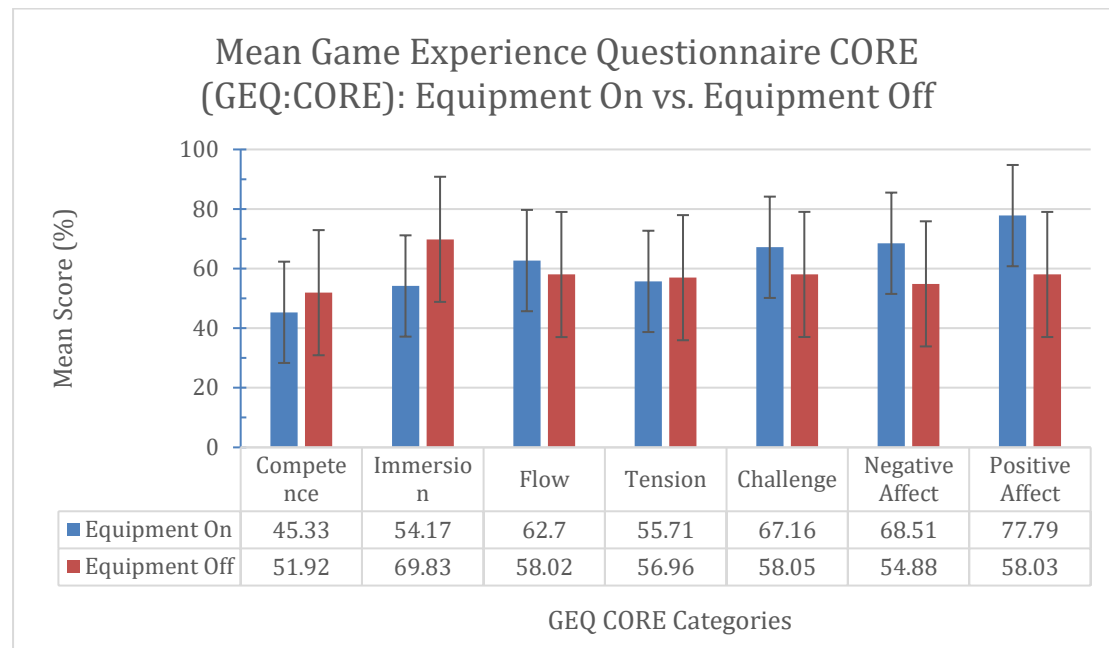


Figure 73: Chart illustrating mean differences between Equipment On and Equipment Off conditions for each of the GEQ CORE components.

Paired-sample t-tests comparing mean differences for each of the GEQ CORE components between Equipment On and Equipment Off conditions revealed that the presence of equipment elicited a range of influences upon subjective game experience (see also Table 55 below):

- **Similar Competence:** Presence of equipment elicited a small strength (*Cohen's d* = .23) statistically **insignificant** mean decrease of -6.58 (95% CI, -15.13 to 1.96) in subjective Competence compared to no equipment being present,  $t(43) = -1.55$ ,  $p = .128$ .
- **Decreased Immersion:** Presence of equipment elicited a medium strength (*Cohen's d* = .45) statistically **significant** mean decrease of -15.66 (95% CI, -26.33 to -4.99) in subjective Immersion compared to no equipment being present,  $t(43) = -2.96$ ,  $p = .005$ .
- **Similar Flow:** Presence of equipment elicited a small strength (*Cohen's d* = .14) statistically **insignificant** mean increase of 4.59 (95% CI, -5.45 to 14.63) in subjective Flow compared to no equipment being present,  $t(43) = .92$ ,  $p = .362$ .

- **Similar Tension:** Presence of equipment elicited a trivial strength (*Cohen's d* = .04) statistically **insignificant** mean decrease of -1.27 (95% CI, -10.86 to 8.33) in subjective Tension compared to no equipment being present,  $t(43) = -.27$ ,  $p = .791$ .
- **Increased Challenge:** Presence of equipment elicited a small strength (*Cohen's d* = .27) statistically **significant** mean increase of 9.12 (95% CI, -18.48 to 1.96) in subjective Challenge compared to no equipment being present,  $t(43) = 1.96$ ,  $p = .043$ .
- **Increased Negative Affect:** Presence of equipment elicited a medium strength (*Cohen's d* = .48) statistically **significant** mean increase of 13.63 (95% CI, -22.37 to 3.15) in subjective Negative Affect compared to no equipment being present,  $t(43) = 3.15$ ,  $p = .003$ .
- **Increased Positive Affect:** Presence of equipment elicited a large strength (*Cohen's d* = .87) statistically **significant** mean increase of 19.76 (95% CI, -26.68 to 5.76) in subjective Positive Affect compared to no equipment being present,  $t(43) = 5.76$ ,  $p < .0005$ .

**Subjective CORE Game Experience Comparisons: Paired Samples Test**

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	EQ Competence - NEQ Competence	-6.58	28.11	4.24	-15.13	1.96	-1.55	43	.128
Pair 2	EQ Immersion - NEQ Immersion	-15.66	35.09	5.29	-26.33	-4.99	-2.96	43	.005
Pair 3	EQ Flow - NEQ Flow	4.59	33.02	4.98	-5.45	14.63	.92	43	.362
Pair 4	EQ Tension - NEQ Tension	-1.27	31.56	4.76	-10.86	8.33	-.27	43	.791
Pair 5	EQ Challenge - NEQ Challenge	9.12	30.79	4.64	-.25	18.48	1.96	43	.043
Pair 6	EQ Negative Affect - NEQ Negative Affect	13.63	28.74	4.33	4.90	22.37	3.15	43	.003
Pair 7	EQ Positive Affect - NEQ Positive Affect	19.76	22.74	3.43	12.85	26.68	5.76	43	< .0005

Table 56: Results of paired-sample t-tests comparing mean subjective GEQ CORE components of game experience for Equipment On against Equipment Off conditions. Mean differences and the significance of these are demonstrated for each of the seven components.

**4.3.1.2. Subjective SOCIAL Game Experience Comparisons**

The results for each of the three components of the Game Experience Questionnaire SOCIAL (GEQ: SOCIAL) were averaged across all participants, separated into the Equipment On and Equipment Off experimental conditions. Survey data is presented as percentage of maximum score for each component. The mean data for both Equipment On and Equipment Off conditions are presented in Table 57 and Figure 74 below.

**Game Experience Questionnaire SOCIAL  
(GEQ:SOCIAL) Survey Results:**

		Equip. On	Equip. Off
Social Empathy	Mean	59.81	69.28
	Std. Deviation	28.12	20.03
Social Negative Feelings	Mean	62.30	55.75
	Std. Deviation	24.57	25.28
Social Behavioural Involvement	Mean	57.25	65.54
	Std. Deviation	20.98	22.78

Table 57: Descriptive statistics for Game Experience Questionnaire SOCIAL components, averaged across all participants, and separated into Equipment On and Equipment Off conditions. Data in % of maximum score.

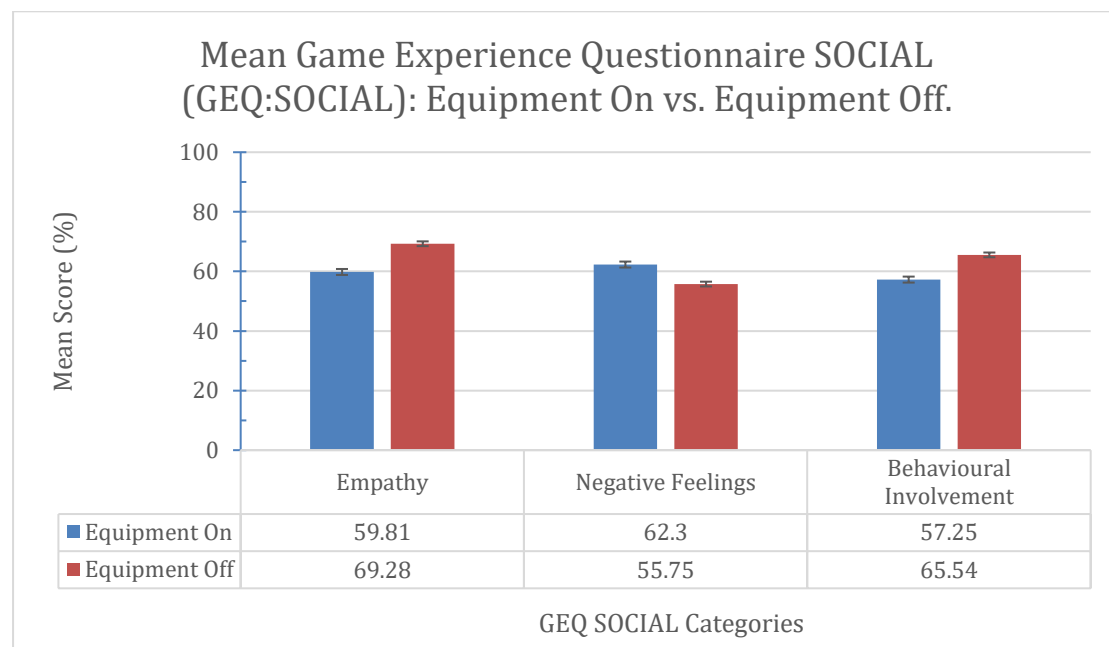


Figure 74: Chart illustrating mean differences between Equipment On and Equipment Off conditions for each of the GEQ SOCIAL components.

Paired-sample t-tests comparing mean differences for each of the GEQ SOCIAL components between Equipment On and Equipment Off conditions revealed that the presence of equipment elicited a range of influences upon subjective social game experience (see also Table 58 below):

- **Decreased Empathy for opponent:** Presence of equipment elicited a small strength (*Cohen's d* = .30) statistically **significant** mean decrease of -9.48

(95% CI, -19.19 to .24) in subjective Empathy for opponent player compared to no equipment being present,  $t(43) = -1.97$ ,  $p = .041$ .

- **Similar Negative Feelings towards opponent:** Presence of equipment elicited a small strength ( $Cohen's d = .18$ ) statistically **insignificant** mean increase of 6.54 (95% CI, -4.26 to 17.34) in subjective Negative Feelings for opponent player compared to no equipment being present,  $t(43) = 1.22$ ,  $p = .228$ .
- **Decreased Behavioural Involvement with opponent:** Presence of equipment elicited a small strength ( $Cohen's d = .26$ ) statistically **significant** mean decrease of 8.29 (95% CI, -18.07 to 1.49) in subjective Behavioural Involvement with opponent player compared to no equipment being present,  $t(43) = -1.71$ ,  $p = .045$ .

#### Subjective SOCIAL Game Experience Comparisons: Paired Samples Test

		Paired Differences							
			Std.	Std.	95% Confidence				
		Mean	Deviation	Error	Interval of the		t	df	Sig. (2-
				Mean	Lower	Upper			tailed)
Pair 1	EQ Social Empathy - NEQ Social Empathy	-9.48	31.95	4.82	-19.19	.24	-1.97	43	.041
Pair 2	EQ Social Negative Feelings - NEQ Social Negative Feelings	6.54	35.52	5.36	-4.26	17.34	1.22	43	.228
Pair 3	EQ Social Behavioural - NEQ Social Behavioural	-8.29	32.17	4.85	-18.07	1.49	-1.71	43	.045

Table 58: Results of paired-sample t-tests comparing mean subjective GEQ SOCIAL components of game experience for Equipment On against Equipment Off conditions. Mean differences and the significance of these are demonstrated for each of the three components.

#### 4.3.1.3. Subjective Arousal and Valence: Equipment On vs. Equipment Off

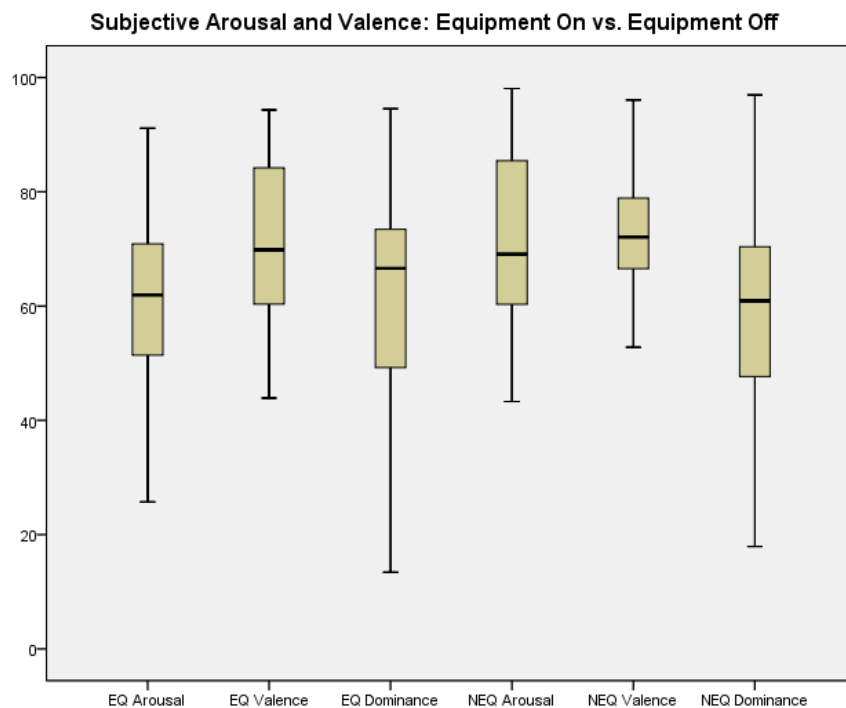
Data on subjective arousal and valence was collected through the Self-Assessment Manikin (SAM), which gives data in percentages of maximum score for three categories of affective experiences (Arousal, Pleasure, and Dominance). The SAM measure of Pleasure is analogous with valence (Nacke, Nacke and Lindley, 2009), and so will be referred to as valence from here on in.

A Shapiro-Wilk test for normality was conducted for all SAM data, with all passing and therefore being normally distributed (see Table 59 below). Additionally, no outliers were evident in any of the data upon inspection of a box plot for values greater than 1.5 box-lengths from the edge of the box (see Figure 75 below for sample).

**SAM Survey Data: Tests of Normality**

	Statistic	Shapiro-Wilk df	Sig.
EQ Arousal	.986	44	.858
EQ Valence	.953	44	.170
EQ Dominance	.912	44	.130
NEQ Valence	.953	44	.173
NEQ Pleasure	.969	44	.285
NEQ Dominance	.988	44	.911

Table 59: Results of Shapiro-Wilk test for normality, conducted for the Self-Assessment Manikin components for both Equipment On and Equipment Off experimental conditions. All data was found to be normally distributed.



*Figure 75: Box plot illustrating the distribution of subjective SAM components for both Equipment On and Equipment Off conditions. Note that no outliers are identified.*

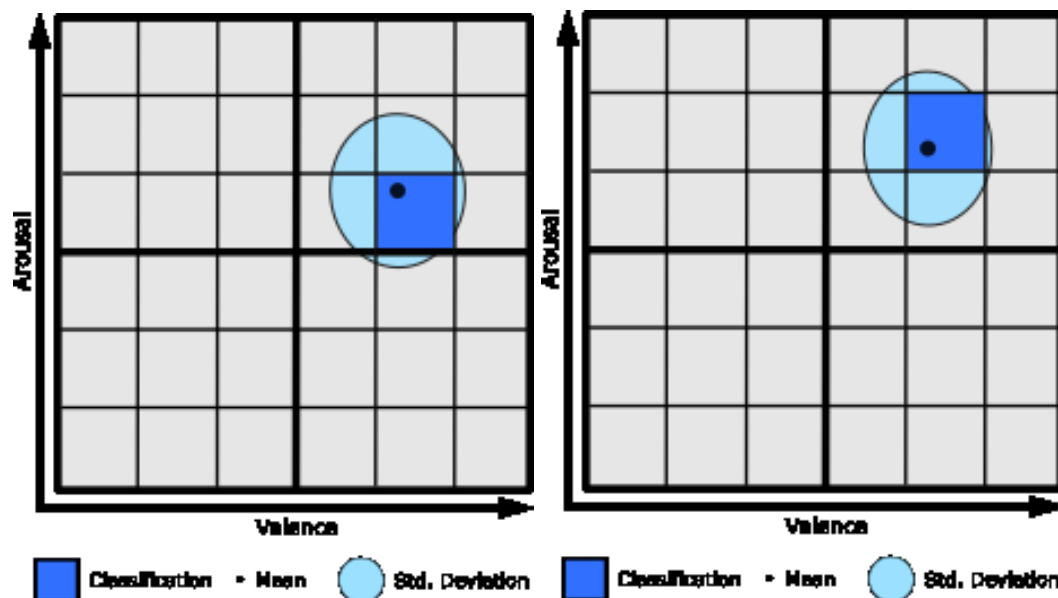
The subjective arousal and valence scores were used to map arousal and valence onto the Affect Grid, illustrated in Figure 76 and Figure 77 below. The Affect Grids demonstrate that playing with Equipment On was perceived to be mid-level arousing with mid-level positive valence, while playing with Equipment Off was perceived to be higher-level arousing with the same mid-level positive valence.



### The Influence of Physiological Equipment on Affect

Left: Equipment On.

Right: Equipment Off.



Figures 76 and 77: Affect Grids modelling the mean subjective arousal and valence scores for Equipment On (Left Affect Grid) and Equipment Off (Right Affect Grid). Valence stays similar for both conditions, however arousal appeared to increase when Equipment was not used.

The results for each of the three components of the Self-Assessment Manikin (SAM) were averaged across all participants, separated into the Equipment On and Equipment Off experimental conditions. Survey data is presented as percentage of maximum score for each component. The mean data for both Equipment On and Equipment Off conditions are presented in Table 60 and Figure 78 below.

#### Self-Assessment Manikin (SAM) Survey Results:

		Equip. On	Equip. Off
Arousal	Mean	61.24	71.39
	Std. Deviation	15.43	15.50
Pleasure	Mean	71.77	72.86
	Std. Deviation	13.98	10.09

<b>Dominance</b>	Mean	59.16	59.37
	Std. Deviation	18.12	18.12

Table 60: Descriptive statistics for SAM components, averaged across all participants, and separated into Equipment On and Equipment Off conditions. Data in % of maximum score.

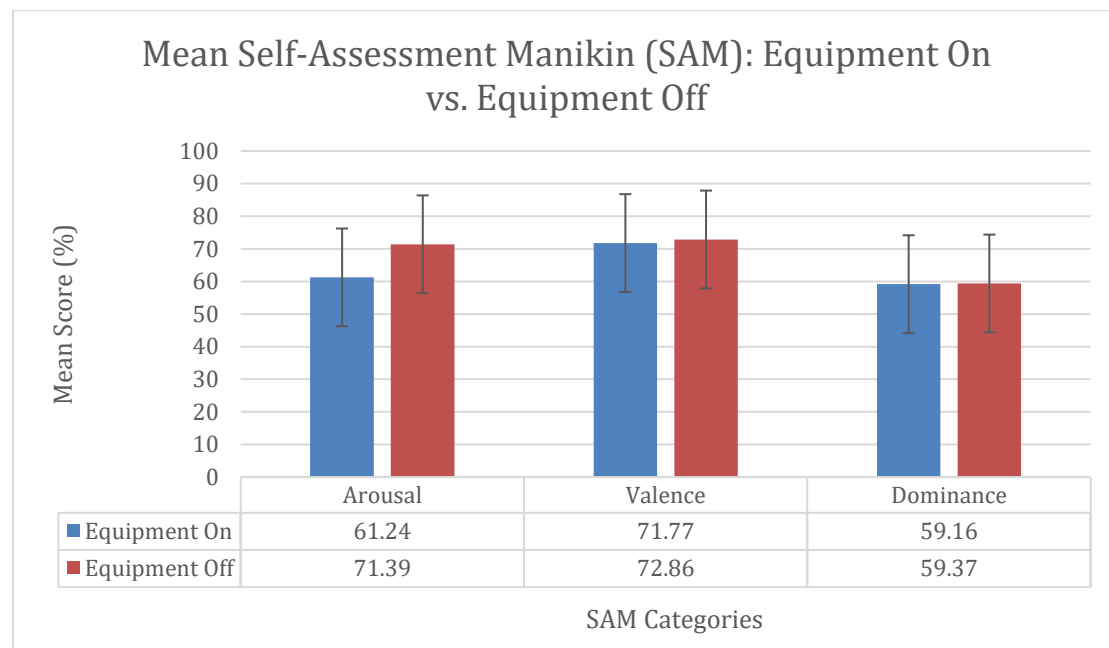


Figure 78: Chart illustrating mean differences between Equipment On and Equipment Off conditions for each of the SAM components.

Paired-sample t-tests comparing mean differences for each of the SAM components between Equipment On and Equipment Off conditions revealed that the presence of equipment elicited a range of influences upon subjective affective experience (see also Table 61 below):

- **Decreased Arousal:** Presence of equipment elicited a medium strength (*Cohen's d* = .48) statistically **significant** mean decrease of -10.15 (95% CI, -16.61 to -3.69) in subjective Arousal compared to no equipment being present,  $t(43) = -3.17$ ,  $p = .003$ .
- **Similar Valence:** Presence of equipment elicited a trivial strength (*Cohen's d* = .06) statistically **insignificant** mean decrease of -1.09 (95% CI, -6.42 to 4.24)

in subjective Valence compared to no equipment being present,  $t(43) = -.41$ ,  $p = .683$ .

- **Similar Dominance:** Presence of equipment elicited a trivial strength (*Cohen's*  $d = .007$ ) statistically **insignificant** mean decrease of  $-.21$  (95% CI,  $-8.51$  to  $8.09$ ) in subjective Dominance compared to no equipment being present,  $t(43) = -8.51$ ,  $p = .96$ .

**Subjective Affective Experience Comparisons: Paired Samples Test**

		Paired Differences							
					95% Confidence Interval of the Difference				Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	
Pair 1	EQ Arousal - NEQ Arousal	-10.15	21.25	3.20	-16.61	-3.69	-3.16	43	.003
Pair 2	EQ Valence - NEQ Valence	-1.09	17.53	2.64	-6.42	4.24	-.41	43	.683
Pair 3	EQ Dominance - NEQ Dominance	-.21	27.31	4.12	-8.51	8.09	-.05	43	.960

Table 62: Results of paired-sample t-tests comparing mean subjective SAM components of affective experience for Equipment On against Equipment Off conditions. Mean differences and the significance of these are demonstrated for each of the three components.

#### 4.3.1.4. Observed Facial Expressions: Equipment On vs. Equipment Off

Observed facial expressions are reported in mean cumulative duration, in seconds, of either positive or negative expressions being observed (see Section 3.9 for further details). As illustrated in Table 63 below, the mean cumulative duration of positive facial expressions was 23.39 seconds ( $\pm 5.37$  seconds) when equipment was present and 45.6 seconds ( $\pm 9.06$  seconds) when equipment was not present.

In contrast, the mean cumulative duration of negative facial expressions was 12.61 seconds ( $\pm 5.13$  seconds) when equipment was present and 23.92 seconds ( $\pm 9.07$  seconds) when equipment was not present. A Shapiro-Wilk test for normality of distribution was conducted for both positive facial expressions and negative facial

expressions, both passing with  $p = .186$  and  $p = .417$  respectively. The data was inspected for outliers through inspection of a box plot, with no outliers identified.

**Mean Cumulative Observed Facial Expressions: Descriptive Statistics**

	N	Mean	Std. Deviation
EQ Negative Expressions (seconds)	44	12.61	5.13
NEQ Negative Expressions (seconds)	44	23.92	9.07
EQ Positive Expressions (seconds)	44	23.39	5.37
NEQ Positive Expressions (seconds)	44	45.60	9.06

Table 63: Descriptive statistics for observed facial expressions, averaged across all participants, and separated into Positive and Negative for both Equipment On and Equipment Off conditions. Data in seconds of expression duration.

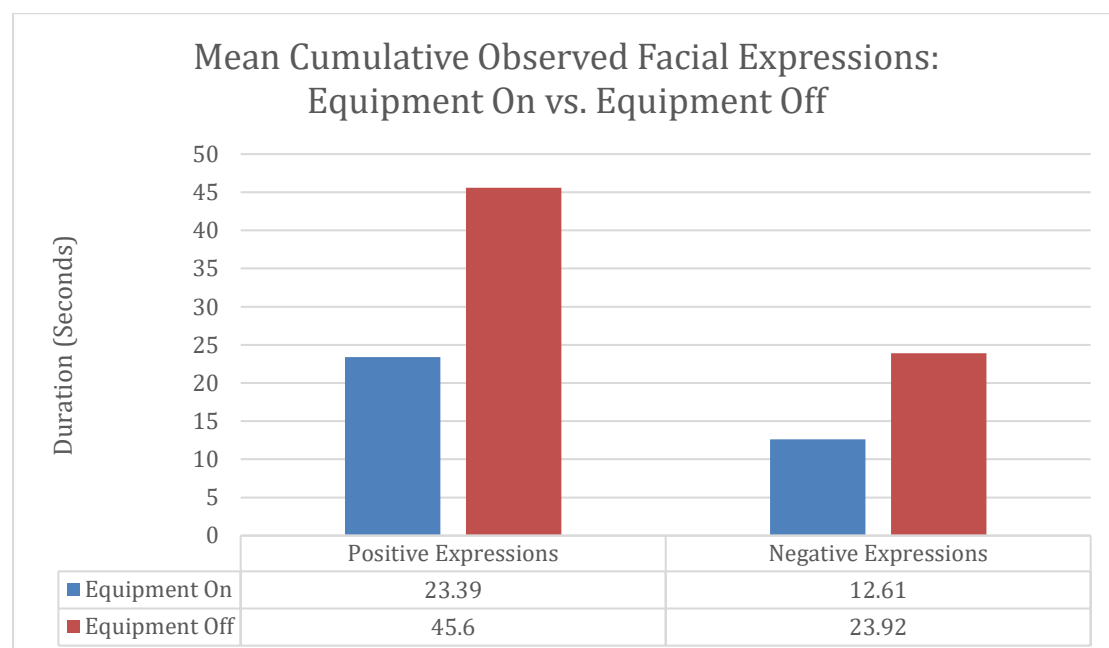


Figure 79: Chart illustrating mean cumulative differences between Equipment On and Equipment Off conditions for both Positive and Negative Expressions.

Paired-sample t-tests comparing mean differences for both Positive and Negative Observed Facial Expressions between Equipment On and Equipment Off conditions

revealed that the presence of equipment elicited a range of influences upon observed facial expressions (see also Table 64 below):

- **Decreased Negative Facial Expressions:** Presence of equipment elicited a large strength (*Cohen's d* = .93) statistically **significant** mean decrease of -11.32 (95% CI, -15.03 to -7.61) in observed Negative Facial Expressions compared to no equipment being present,  $t(43) = -6.15, p < .0005$ .
- **Decreased Positive Facial Expressions:** Presence of equipment elicited a large strength (*Cohen's d* = 2.39) statistically **significant** mean decrease of -22.21 (95% CI, -25.02 to -19.39) in observed Positive Facial Expressions compared to no equipment being present,  $t(43) = -15.89, p < .0005$ .

**Observed Facial Expression Comparisons: Paired Samples Test**

		Paired Differences							
				Std.	95% Confidence				
		Mean	Std.	Error	Interval of the		t	df	Sig. (2-
			Deviation	Mean	Lower	Upper			tailed)
Pair 1	EQ Negative Expressions - NEQ Negative Expressions	-11.32	12.20	1.84	-15.03	-7.61	-6.15	43	< .0005
Pair 2	EQ Positive Expressions - NEQ Positive Expressions	-22.21	9.27	1.40	-25.02	-19.39	-15.89	43	< .0005

Table 64: Results of paired-sample t-tests comparing mean observed Positive and Negative Facial Expressions for Equipment On against Equipment Off conditions. Mean differences and the significance of these are demonstrated for each of the three components.

#### 4.3.2. Comparing Objective and Subjective Measures

Data on subjective valence and arousal was collected through the Self-Assessment Manikin (SAM) survey instrument, while data on objective valence and arousal was collected through physiological signals and modelled using a fuzzy logic system (see

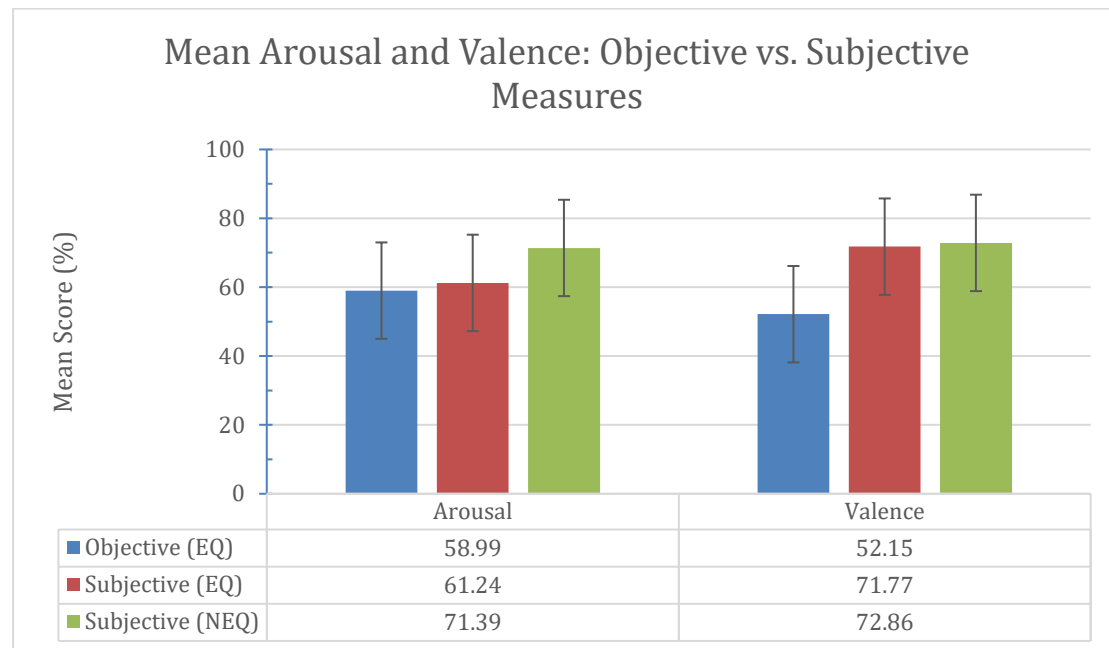
Section 3.8). While the influence that the use of psychophysiological equipment has upon subjective measures of game experience and affective experiences has been established within previous sections of this results chapter, ostensibly both the SAM and fuzzy logic systems measure the same aspects of player experiences, and so it remains to examine for relationship between these subjective and objective measures.

Shapiro-Wilk tests for normality have previously been conducted for all data within this section and, as reported in previous sections, all data was found to be normally distributed. Additionally, no outliers were evident in any of the data upon inspection of a box plot for values greater than 1.5 box-lengths from the edge of the box.

### Comparison between Subjective and Objective: Measures of Valence and Arousal:

		Equip. On	Equip. Off
Subjective Arousal	Mean	61.24	71.39
	Std. Deviation	15.43	15.50
Subjective Valence	Mean	71.77	72.86
	Std. Deviation	13.98	10.09
Objective Arousal	Mean	58.99	N/A
	Std. Deviation	12.99	N/A
Objective Valence	Mean	52.15	N/A
	Std. Deviation	18.72	N/A

*Table 65: Descriptive statistics comparing Subjective and Objective measures of Valence and Arousal, averaged across all participants, and separated into Equipment On and Equipment Off conditions. Data in % of maximum score.*



*Figure 80: Chart illustrating mean differences in Arousal and Valence under Equipment On and Equipment Off conditions, for both Objective and Subjective measures.*

#### **4.3.2.1. Correlations between Subjective and Objective Measures**

Pearson's tests of correlation were run to assess the relationship between Subjective Arousal and Objective Arousal, as well as between Subjective Valence and Objective Valence. Preliminary analyses for both sets of relationships showed that all variables were normally distributed, as assessed by Shapiro-Wilk's test ( $p > .05$ ), and there were no outliers within the data sets. Scatterplots revealed general linear trends across Subjective and Objective measures for both Arousal and Valence (see Figure 81 and Figure 82 below)

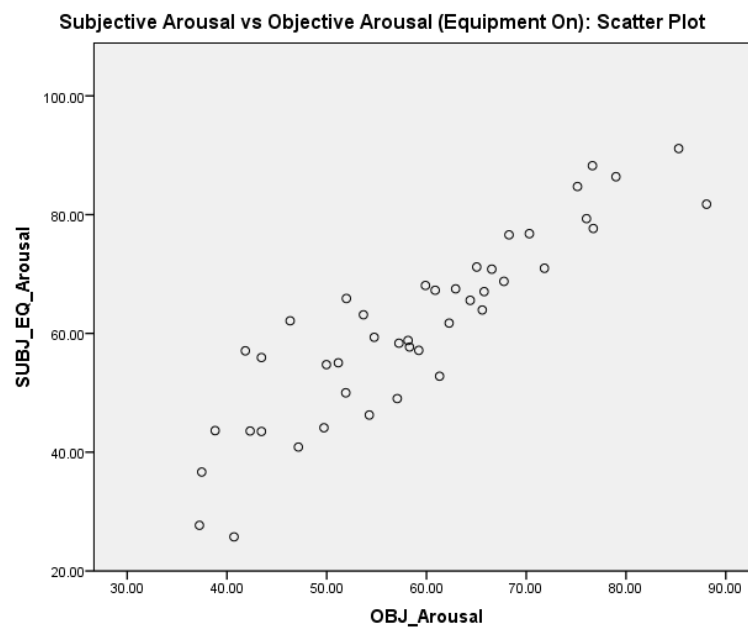


Figure 81: Scatterplot modelling Subjective Arousal against Objective Arousal to investigate the relationship between measures. Note a general linear relationship.

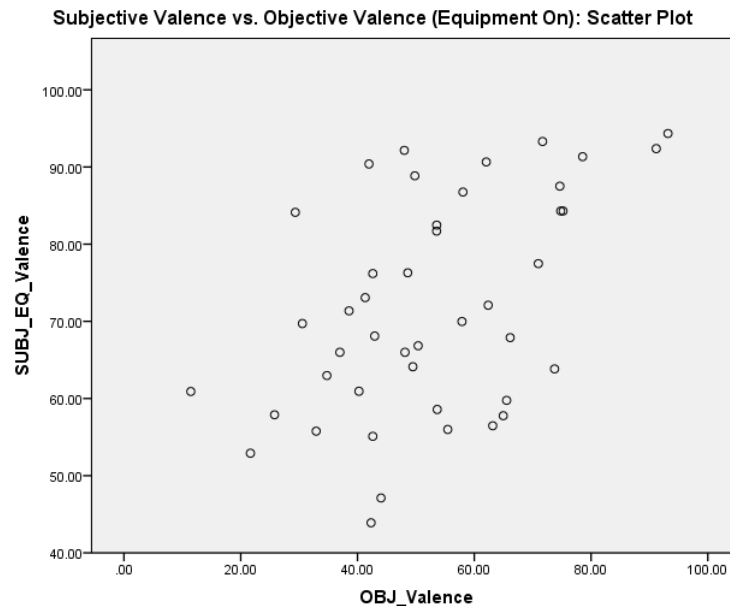


Figure 82: Scatterplot modelling Subjective Valence against Objective Valence to investigate the relationship between measures. Note a general linear relationship.



There was a strong positive correlation between Subjective Arousal and Objective Arousal measures,  $r(44) = .896$ ,  $p < .0005$ , with the Subjective Arousal measure explaining 80% of the Objective Arousal measure (see Table 66 below).

**Subjective Arousal vs Objective Arousal: Pearson's Correlation**

		Subjective Arousal	Objective Arousal
Subjective Arousal	Pearson Correlation	1	<b>.896**</b>
	Sig. (2-tailed)		<b>&lt; .0005</b>
	N	44	44
Objective Arousal	Pearson Correlation	<b>.896**</b>	1
	Sig. (2-tailed)	<b>&lt; .0005</b>	
	N	44	44

\*\* Correlation is significant at the 0.01 level (2-tailed).

Table 66: Pearson's test for correlation between Subjective and Objective measures of Arousal, demonstrating a strong positive relationship.

There was a medium positive correlation between Subjective Valence and Objective Valence measures,  $r(44) = .496$ ,  $p = .001$ , with the Subjective Valence measure explaining 25% of the Objective Valence measure (see Table 67 below).

**Subjective Valence vs Objective Valence: Pearson's Correlation**

		Subjective Valence	Objective Valence
Subjective Valence	Pearson Correlation	1	<b>.496**</b>
	Sig. (2-tailed)		<b>.001</b>
	N	44	44
Objective Valence	Pearson Correlation	<b>.496**</b>	1
	Sig. (2-tailed)	<b>.001</b>	
	N	44	44

\*\* Correlation is significant at the 0.01 level (2-tailed).

Table 67: Pearson's test for correlation between Subjective and Objective measures of Valence, demonstrating a medium positive relationship.

### 4.3.3. Summary of Results for Research Question 2

The previous sections have presented the results of descriptive and inferential statistical analysis into the subjective game and affective experiences of participants as measured at the conclusion of game play, and demonstrates how the presence of psychophysiological equipment influenced these subjective experiences in a variety of ways. Research Question 2 sought to explore influences that the use of psychophysiological instruments might have upon the subjective game play and affective experiences of participants playing video games under experimental conditions, and examine the effects that this had upon the objective psychophysiological data obtained for the purposes of addressing RQ1 (see Section 4.2 above).

As evidenced in Figure 83 below, the subjective measures included within this analysis suggest that the use of psychophysiological instruments elicit a range of influences upon the subjective game experience and subjective affective state of participants, the implications for which are discussed in Chapter 5: Discussions.

Additionally, this section presented evidence that despite the influence that the objective psychophysiological instruments appeared to have upon subjective game play experiences, the subjective measures of arousal and valence positively and significantly correlate with the objective measures of arousal and valence. The implications for this are discussed in Section 5.4.

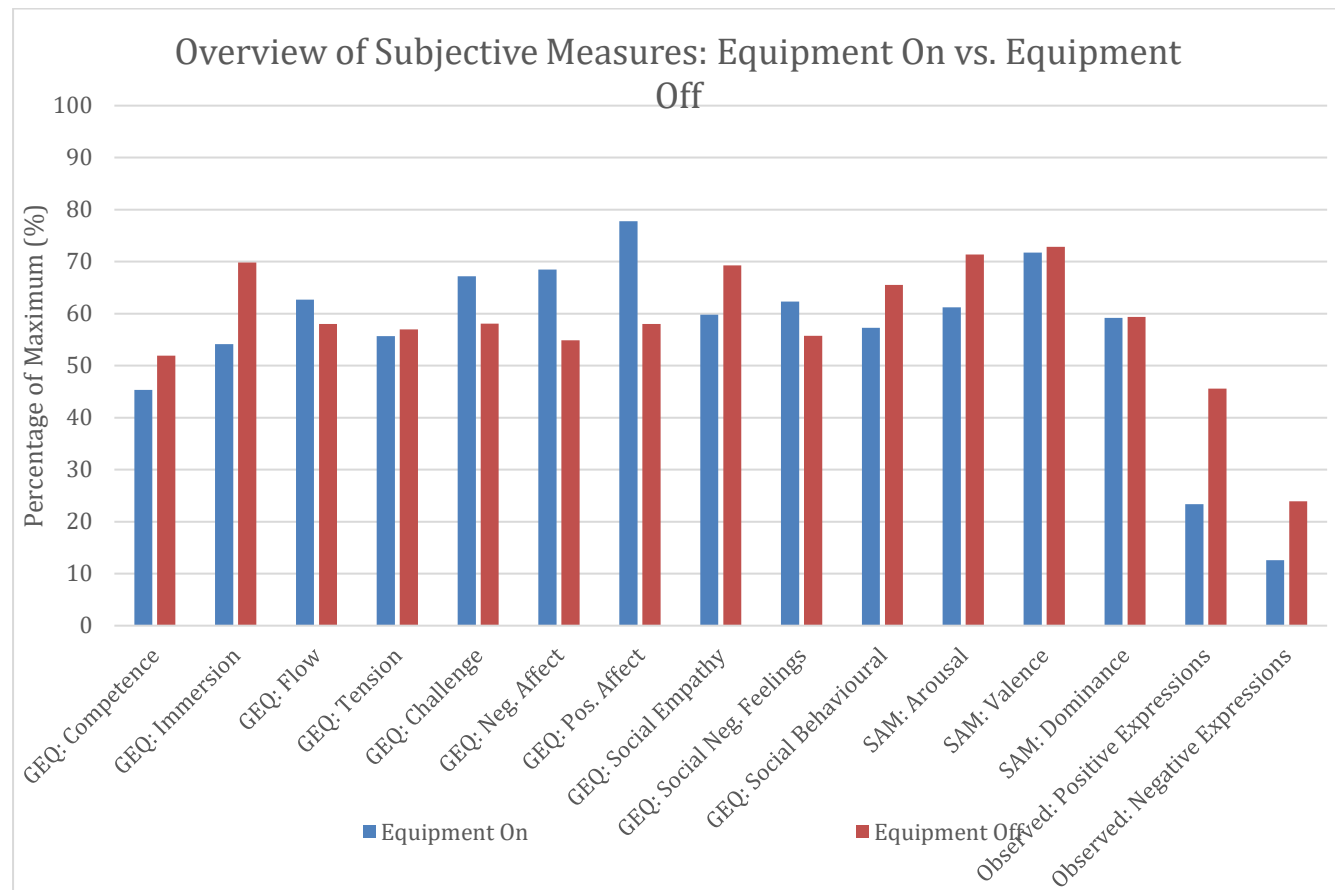


Figure 83: Overview of the subjective measures analysed within the results section of this project for both Equipment On and Equipment Off conditions.

All data in mean % of maximum score except for Observed Positive and Negative Expressions, which are in mean seconds of expression duration. In-depth analysis, including effect size, strength, and significance of differences is presented in the previous sections.

#### 4.4. Summary of Results Chapter

This results chapter has presented the results of multiple sets of analyses that demonstrate the influences that ten video game events were found to have upon affective responses. Additionally, analyses were presented that demonstrate the influence that the use of psychophysiological instruments had upon perceived game experiences, despite significant positive correlations between subjective and objective measures of affect and valence. These results, and their respective interpretations and implications, are discussed in the following Discussions chapter.

## Chapter 5: Discussions

### 5.1. Layout of Chapter / Structure

This chapter has been developed to discuss the implications of the research findings that have been presented so far. Specifically, this chapter will discuss the findings with reference to the two research questions passed earlier in this thesis, including a discussion of relevant literature that assists in the interpretation of the results, and places them within the contexts of the existing body of theory.

### 5.2. Introduction

This study explored the use of psychophysiology as a measurement tool for the capture and assessment of nuanced affective experiences during video game play, with an emphasis on understanding how events occurring during game play elicit specific affective responses that can be measured and understood separately from the overall or average game experience.

Participants were continuously monitored for Heart Rate (HR), Galvanic Skin Response (GSR), and facial Electromyographic (EMG) changes during video game play, which were classified into arousal and valence scores using a fuzzy logic system, and then further classified into membership of five categories of affective experiences using a second fuzzy logic system (see Section 3.8.2) to facilitate objective analyses of the affective experiences of participants. Additionally, the Game Experience Questionnaire (GEQ) and Self-Assessment Manikin (SAM) surveys were used as measures of perceived game experience, and to detect changes in game experience related to the presence, or absence, of the psychophysiological equipment used to obtain objective data.

It is important to note that although a series of game events were identified and explored within this study, the results of each event are not intended to be generalizable to other video games. This is because games typically use a unique set of mechanics, and therefore events, to create a game experience that may differ

substantially from the experiences that may be elicited from other games. Rather, it is the objective of this study to explore the ability of game events to be identified and explored in terms of their respective affective responses using objective psychophysiological measures. The results of game events are therefore discussed in terms of what can be learned about affective responses to differing game, rather than how these specific game events may be relevant and generalizable to video games other than Tetris Party Deluxe, the game used in this study.

Two sets of analyses were carried out, one for each Research Question previously presented (see Sections 4.2 and 4.3 respectively).

Guided by objective measures for Arousal and Valence, which were classified into five categories of emotions: Fun, Excitement, Boredom, Frustration, and Challenge, the analysis of data pertinent to Research Question 1 (RQ1) identified ten events that occurred during video game play, and explored the affective responses that each of these elicited. To this end, a series of statistical tests were undertaken for each of the game events, with results indicating significant effects upon elicited affective responses. These results are discussed within the following sections of this chapter.

The second set of analyses, undertaken for Research Question 2 (RQ2), explored the interactions between perceived game experiences when the use of psychophysiological equipment was experimentally manipulated. The analysis of data pertinent to RQ2 identified several significant effects that such equipment appears to have upon perceived game experiences, the implications of which are discussed within subsequent sections of this chapter.

### 5.3. Research Question 1

#### ***Research Question 1:***

*How do events occurring during video game play influence event-related affective responses?*

Research Question 1 aimed to explore the influence that certain video game events can have upon affective responses as experienced during real-time game play. The unit of analysis was therefore observed game events, analysing the influence that these had upon affective responses during game play. This required use of modelled arousal and valence (AV space), as well as the subsequently modelled emotional categories of fun, excitement, boredom, frustration, and challenge (see Section 3.8 for further details on the modelling process).

The modelled AV space and emotional categories provide real-time objective insights into affective states and responses to ten identified game events, which were observed and recorded as part of the data processing prior to analysis (see Section 3.9.3).

In addition to the mean game play experience, the ten events included within this study were:

- **Mean game play:** Objective measures averaged across play sessions for the aggregate participant population. Events are compared against mean game play to establish differences between event-related responses and general experiences.
- 11. **Game Speed Increase:** Events where the game speed periodically increased in a cumulative mechanic, resulting in players continually having to respond more quickly.
- 12. **Non-Useful Tetriminos:** Events where players received Tetriminos that were not useful for their current game board composition, and thereby hindered performance.
- 13. **Correct Placement:** Events where players placed Tetriminos where they had intended, in a location that either improved performance, or at least did not hinder it.

14. **Incorrect Placement:** Events where players placed Tetriminos that either hindered performance, or where players responded in a way to indicate that the Tetriminos were misplaced.
15. **Multiple Line Clear:** Events where players cleared multiple lines of Tetriminos from the game board simultaneously.
16. **Player Win:** Events where players won the game against their opponent.
17. **Player Loss:** Events where players lost the game against their opponent.
18. **Opponent Correct Placement:** Events where the player observed or commented on their opponent placing Tetriminos in a location that improved their respective performance.
19. **Opponent Incorrect Placement:** Events where the player observed or commented on their opponent placing Tetriminos in a location that hindered their respective performance.
20. **Opponent Multiple Line Clear:** Events where the player observed or commented on their opponent clearing multiple lines of Tetriminos from the game board simultaneously.

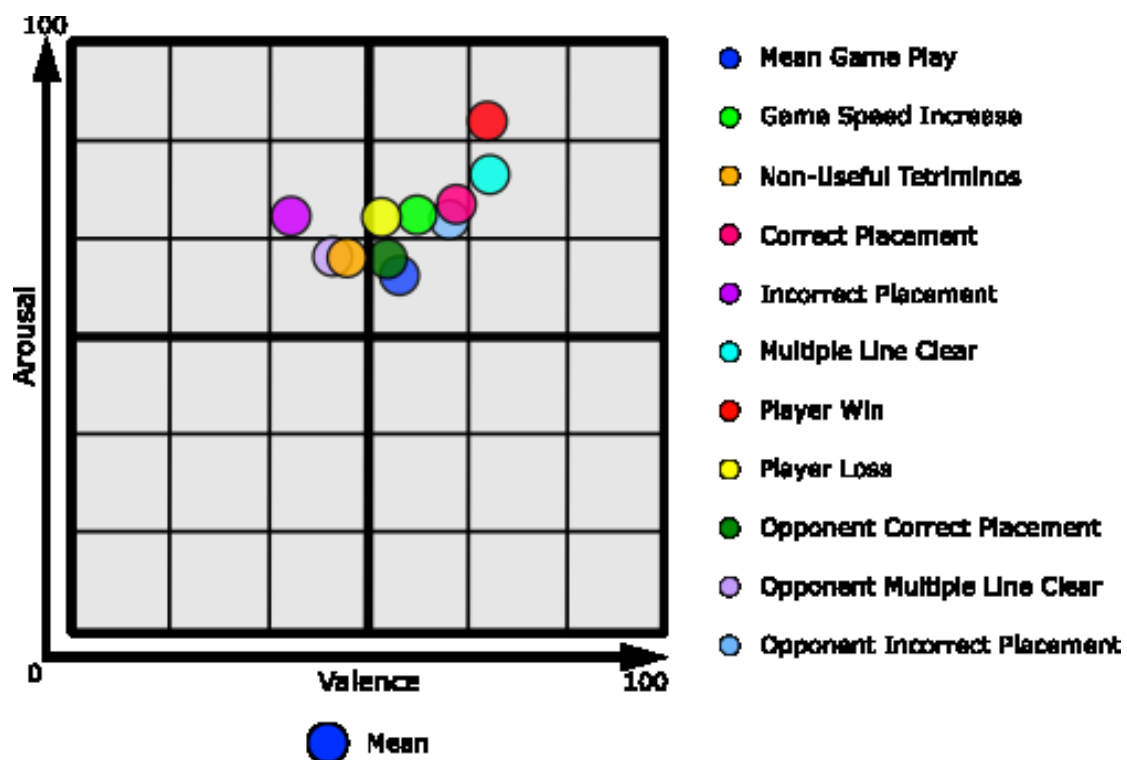
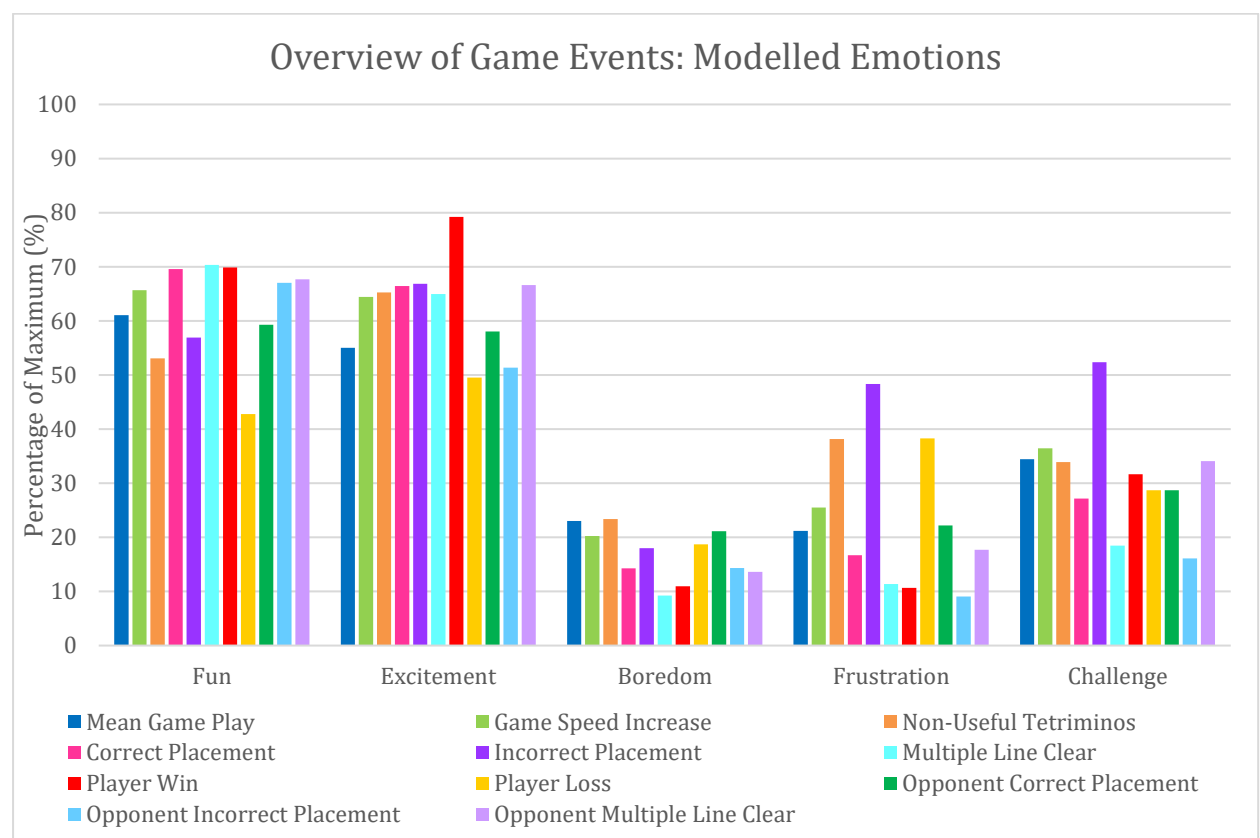


Figure 70 (repeated): Overview of the modelled AV space for each of the game events analysed within the results section of this project. In-depth analysis,

*including effect size, strength, and significance of differences is presented in the Results chapter.*

As illustrated in Figure 70 above and Figure 71 below, the ten game events differed from mean game play across arousal, valence, and the five emotions included within this study. Interpretations and respective implications of these affective responses to observed game events are discussed within the following sections.



*Figure 71(repeated): Overview of the modelled Emotional categories for each of the game events analysed within the results section of this project. In-depth analysis, including effect size, strength, and significance of differences is presented in the Results chapter.*

### 5.3.1. Mean Affective Game Experiences

The primary aim of RQ1 was to explore affective responses elicited through game events experienced by participants during video game play. This was facilitated



through the mapping of objective physiological signals to a modified circumplex affect grid, modelling the Arousal Valence space (AV space) using a fuzzy logic system (see Section 3.8.1 for details). The modelled AV space provided a measure of the overall, or mean, affective game experience for participants, against which the identified game events are compared for the purposes of exploring the influences that each had upon event-response affect.

The modelled AV space was subsequently used to model five categories of emotion through a second fuzzy logic system (see Section 3.8.2 for details), providing a breakdown of more explicit emotions as they relate to video game play and the events occurring therein.

The circumplex model of affect posits that Valence and Arousal subserve all emotional experiences and act as a foundation upon which affective experiences can be interpreted and understood (Posner, Russell, and Peterson, 2005). Illustrated in Figure 86 below, emotions can be mapped to a two-dimensional representation of the circumplex model depending upon their respective influence upon Arousal and Valence (Posner, Russell, and Peterson, 2005).

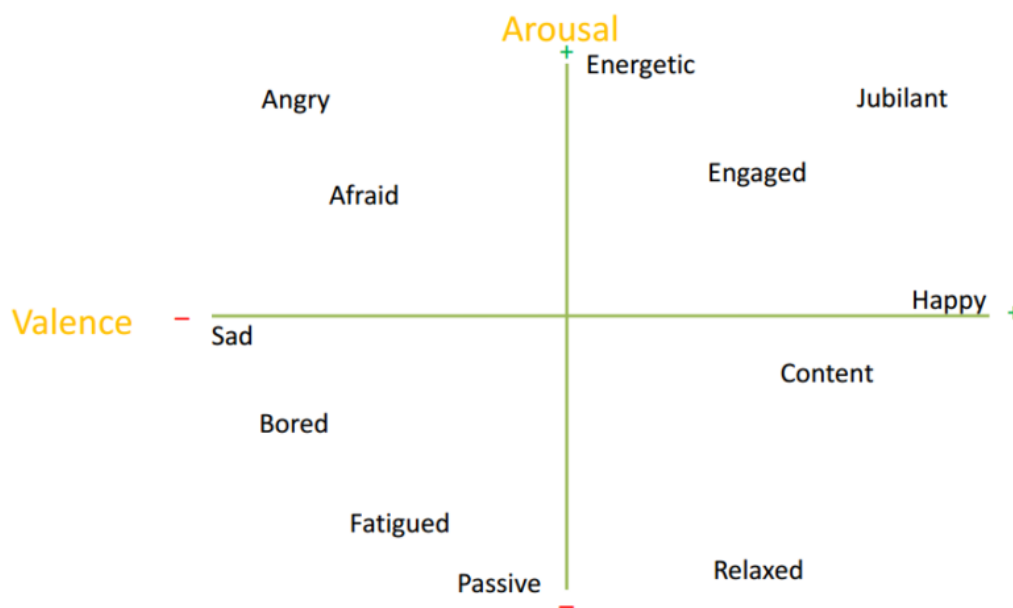


Figure 13 (repeated): A sample Circumplex model of affect representing the two-dimensional space of Arousal and Valence, and the emotions associated with different coordinates within the model (Ambinder, 2011).

#### 5.3.1.1. Mean Valence

Valence is understood on a continuum between absolute negative (0%), such as being sad or depressed, and absolute high (100%), such as being happy (Cacioppo and Tassinary, 1990; Cacioppo, Tassinary and Berntson, 2007). This means that a neutral score of 50% might occur when an individual is experiencing emotions or moods that are neither negative nor positive in nature.

The mean modelled AV space for this study revealed that the overall affective game experience was only a little over neutral valence (50%), with a mean value of 52.15% (+/- 18.72%). This result therefore suggests that the valence of overall game play tended to be experienced as slightly pleasurable or happy state (Cacioppo, Tassinary and Berntson, 2007; Cowley *et al.*, 2016). Within the context of game events examined the mean valence of game play sits within the middle of valence results elicited through game events, with five game events eliciting lower valence relative to mean game play, and five events eliciting higher valence relative to mean game play (see Figure 70).

The mean valence response being placed in the middle of game event responses and slightly positive supports the argument that it was the game events themselves that influenced the valence of the mean game play experience, with three game events leading to negative valence (Non-Useful Tetriminos, Incorrect Placement, and Opponent Multiple Line Clear), while the other seven game events lead to positive valence over the course of the game (refer to Figure 71 above). This further suggests that the game experience between events was relatively neutral in respect to their influence upon valence, as differences in mean valence appears to be explained by the valence responses of observed game events.

Overall, given that valence was recorded physiologically, the valence results appear to be relatively low compared to what might be expected from video game play, with the mean valence being only a little above neutral. Given that video games are typically undertaken as a form of entertainment or as a recreational leisure activity, they may be expected to elicit higher levels of valence, which does not appear to have been the case here. One possibility for this is that the game used, Tetris Party Deluxe, may not have been as engaging or exciting as other modern games, potentially inhibiting the experience of positively valenced affective responses.

Another possible explanation for the relatively low valence of mean game play, which is explored in greater depth in the following sections, is that the equipment used for objective measures of valence (facial expressions as recorded using facial EMG, see Section 3.4.2) appear to significantly inhibit elicitation facial expressions during game play. When objective psychophysiological instruments were being used to measure physiological signals, mean cumulative facial expressions were observed as being 23.39 (+/- 5.37) seconds for positive expressions and 12.61 (+/- 5.13) seconds for negative expressions. In contrast, mean cumulative facial expressions were observed as being 45.60 (+/- 9.06) seconds for positive expressions and 23.92 (+/- 9.07) seconds respectively on participants where the psychophysiological equipment was not being used.

As presented within Section 4.2.1 and discussed within Section 5.3.1, these indicate large significant influences upon the expression of both positive and negatively valenced facial expressions, potentially influencing the ability of psychophysiology to appropriately classify valence from the measures.

### **5.3.1.2. Mean Arousal**

Arousal is understood on a continuum between low-activity (0%), such as being passive or relaxed, and high-activity (100%), such as being excited or energetic (Cacioppo, Tassinary and Berntson, 2007; Cowley *et al.*, 2016). This means that a neutral score of 50% might occur when an individual is sitting quietly and thinking about something, and is neither relaxed nor excited. Similarly, with valence, a neutral score of 50% would suggest that emotions being experienced are neither positive nor negative in nature.

The mean modelled AV space for this study revealed that the overall affective game experience elicited a mean arousal that was only a little over neutral at 58.77% (+/- 14.54%). This result therefore suggests that the arousal of overall game play tended to be experienced as a slightly exciting state (Posner, Russell and Peterson, 2005; Cacioppo, Tassinary and Berntson, 2007; Cowley *et al.*, 2016). Within the context of game events examined the mean arousal of game play had the lowest arousal result elicited through game play, with all 10 game events eliciting higher arousal compared to mean game play, albeit some not significantly or with small effects (see Figure 84).

The mean arousal response being lower than game event responses makes it likely that game events were responsible for the majority, if not all, increases in arousal, and that game play between events was substantially less arousing than the experience of the observed game events. This suggests that mean arousal itself is likely a product of event-related arousal, with differences between mean arousal

and event-response arousal appearing to be explained by the arousal responses to observed game events.

Discussed further in Section 5.3.2 below, this inherently makes sense, as the identification and inclusion of game events was relatively holistic in terms of the game events and interactions included. The implication being that the time participants spent between game events may have been relatively passive and waiting for the next event or interaction to occur, which would reduce their level of arousal.

### **5.3.1.3. Summary of Circumplex Model of Affect**

The mean arousal and valence of game play recorded in this study revealed that, compared with observed game events experienced during game play, both mean valence and mean arousal were slightly above neutral values. Mean valence was situated at the median of all observed game events, with five events eliciting lower valent responses and five events eliciting higher valent responses. Mean arousal was situated at a lower level than any of the observed game events, with all game events eliciting arousal above the mean.

As the game events included within this study are relatively holistic, capturing the majority of actions and situations that can occur within the video game used, these results appear congruent with expectations. Video games rely upon a series of mechanics and interactions to create an overall experience, and while the gaps between events are necessary to facilitate pace and reflection (Przybylski, Rigby and Ryan, 2010), the lack of interactions within them can make it difficult to assess how players are using this time. At least in the case of Tetris Party Deluxe, these gaps between events are not part of the core experience as they typically consist of waiting for the next event or interaction to occur. However, these gaps are still important to the game play experience, as players likely use them to prepare for their next action.

It therefore appears logical that these gaps have no apparent influence upon valence, as the gaps are relatively short, offering just enough time for players to reflect upon the game board and plan their next move. If the gaps were longer, it is possible that players may have become bored or frustrated, potentially resulting in negatively valenced responses during the gaps between events. In the case of mean arousal, these gaps between events typically do not require any actions on the part of the player, and so it also appears logical that mean arousal would decrease during these times. This is congruent with findings in other forms of multimedia, demonstrating that lack of stimulation for short periods of time will typically result in participants experiencing neutrally valenced, low arousal affective states (Pattyn *et al.*, 2008), which is why participants are asked to sit quietly between experimental sessions (Cacioppo, Tassinary and Berntson, 2007).

### 5.3.2. Event-Related Affective Responses

Following the modelling of physiological data into a modified circumplex affect grid using a fuzzy logic system, a second fuzzy logic system was used to decompose affective experiences into membership at degrees of belonging into a series of five emotions (see Section 3.8.2 for details). This was conducted for the entire time series for each participant, facilitating the examination of affective responses at any point in time, as well as an overall mean experience.

As discussed in Section 3.9, data for each of the five categories of emotions was collected for each observed game event occurring during game play. This event-related data was then aggregated across all participants, to generate a pool of event-related data for the participant population as a whole. Additionally, the complete gameplay data for each of the five emotions was aggregated across the participant population, facilitating comparisons between the mean experience and the event-related experience for each event type, and for each category of emotion.

As illustrated in Figure 87 below, the five categories of emotions modelled through the fuzzy logic system, originally demonstrated by Mandryk *et al.* (2008) are:

- **Fun:** Emotions related to feeling entertained by or enjoyment of experiences. Positively valenced, but may have a wide range of arousal.
- **Excitement:** Emotions related to feeling energetic or enthusiastic towards experiences. Positively valenced, with a high-level of arousal.
- **Boredom:** Emotions related to feeling weary or impatient towards experiences. Negatively valenced, with a low-level of arousal.
- **Frustration:** Emotions related to feeling upset or annoyed towards experiences. Negatively valenced, with a high-level of arousal.
- **Challenge:** Emotions related to feeling strained or tested by experiences. A wide range of valences, with a high level of arousal.

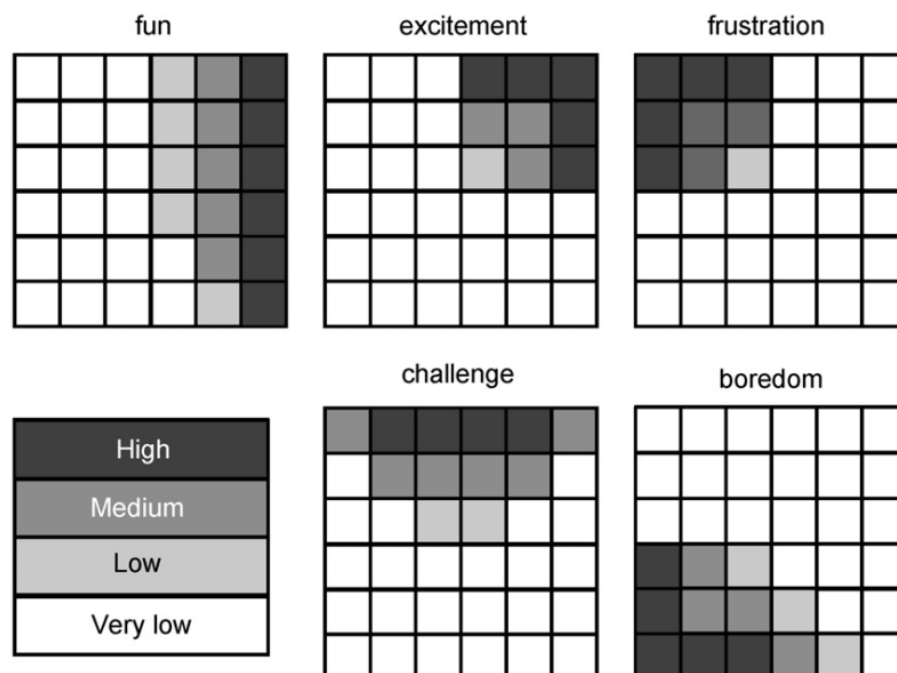


Figure 87 (repeated): Classification of the five categories of emotions based on their respective arousal and valence. From Mandryk *et al.* (2007).

The following sections provide a discussion on each of the ten game events included within this study, including an exploration of the emotions measured for each.

### 5.3.2.1. Game Speed Increase

Increases of Game Speed (coded as “Game Speed Increase” events) occur as players progress in score during Tetris Party Deluxe. The incremental increases in speed causes the Tetriminos to fall more rapidly, reducing the time available for players to react and correctly place them and thereby increasing the difficulty of the game.

As presented in Section 4.2.2, game speed increase events had two main significant effects when compared against mean game play, namely increases in both Excitement and Frustration.

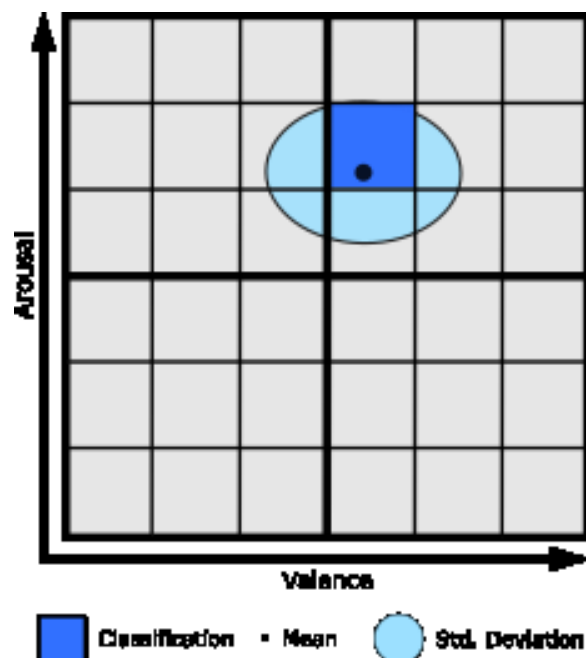
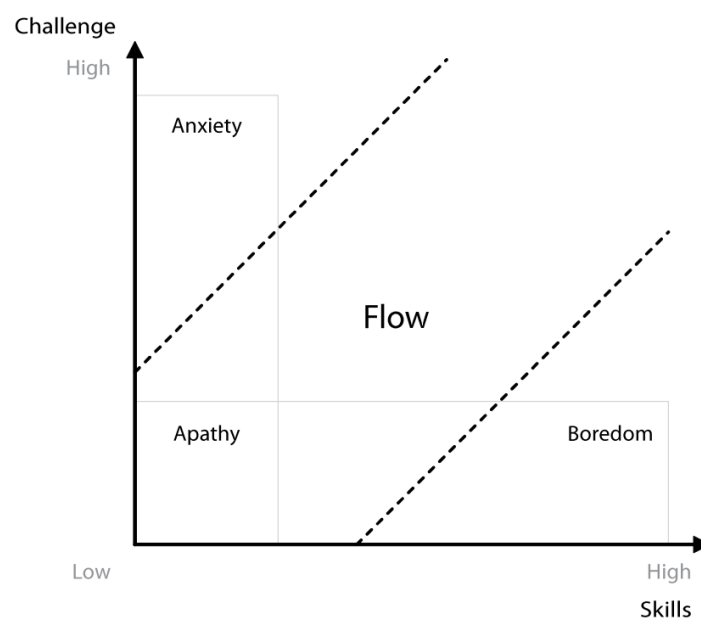


Figure 88: Affect Grid illustrating the mean Arousal and Valence for Game Speed Increase events.

Excitement was found to increase by an average of 9.16% from the mean during game speed increase events, constituting a medium strength effect. As increases in game speed require players to respond more rapidly and thereby increases the difficulty, it is expected that speed increases correspond with increased cognitive engagement. This appears to be congruent with the concept of flow, first introduced by Csíkszentmihályi (1975), which describes flow as an ideal zone of engagement where the skills of the individual are matched to the challenge of the task.



Flow theory has since been adapted and popularised for video game studies, including a recent adaptation by Nacke and Lindlay (2009). Illustrated in Figure 89 below, flow theory describes how players may become apathetic or bored if the challenge is too low, or anxious if the challenge is too high for the skill level of the player. It is common for video games to adapt difficulty over time in an attempt to keep players within an ideal state of flow (Chen, 2007; Nacke and Lindley, 2009), however as Tetris is intended to be a relatively short game per session, the increases in difficulty occur rapidly with the intention of overwhelming the player.



*Figure 89: Nacke and Lindlay's (2009) two-dimensional four-channel model of flow, which they adapted from Csíkszentmihályi (1975) and Ellis, Voelkl, and Morris (1994).*

Flow theory suggests that players will begin to feel anxiety as their skill levels are exceeded during video game play (Nacke and Lindley, 2009; Klasen *et al.*, 2012). Anxiety is an emotional state characterised by negative valence and high arousal (Klimmt *et al.*, 2009; Nacke and Lindley, 2009), and therefore would be characterised through the fuzzy logic system of this study in much the same way that Frustration is characterised. The observed increase in Frustration measured during game speed

increase events (4.34% increase, albeit with a small effect strength) therefore also appears to support the fit with flow theory, as increased frustration appears to align with increased anxiety associated with game speed increasing above the skill level of players.

It is also interesting to note that according to flow theory, an increase of game speed would have been expected to increase the experience of emotions related to Challenge as modelled within this study, due to the increased difficulty in playing the game after game speed increases (Chen, 2007; Nacke and Lindley, 2009; van den Hoogen *et al.*, 2012; Lyons, 2015). However, affective responses related to Challenge only increased by a small amount (1.88%), and did not approach statistical significance.

One possible explanation for this modest impact of the increase of game speed on Challenge could be that while the game speed increases instantly, the real impact of the game speed increase on Challenge could only be felt during the game play over an extended period rather than at the time of the increase occurring. If this is true, then mean challenge for the overall game experience would be elevated instead, but not necessarily at the point of change.

While challenge did not increase during game speed increases, the game speed is a core mechanic that constitutes the primary method through which game difficulty is manipulated, and is likely the primary contributor towards actual challenge within the game. An examination of the GEQ survey results reveals that perceived flow and challenge of the overall game experience were relatively high, at 62.7% of maximum and 67.16% of maximum scores respectively. Compared with the GEQ results of other studies using games that are traditionally considered more immersive, such as the first-person shooter Bioshock examined by Drachen *et al.* (2010), which resulted in GEQ ratings of 67.5% of maximum for perceived flow and 68% of maximum for perceived challenge.

It is therefore considered likely that game speed-increase events were experienced by players as increases in excitement for most the game, at least while the game difficulty increased within their respective skill levels and therefore flow zones. The smaller increases in frustration are therefore likely explained by game speed, and therefore difficulty of the game, as these eventually exceeded the skills of the players, proceeding to elicit frustration or anxiety in accordance with flow theory (Chen, 2007; Nacke and Lindley, 2009).

Given the results, it seems likely that the level of challenge may have been too little for players at the beginning of the experience (resulting in them not being in a state of flow), and too challenging later in the play experience (also resulting in them no longer being in a state of flow). These results, while requiring additional study and examination before drawing firm conclusions, could suggest that flow states may be able to be examined in real-time during game play experiences through the use of psychophysiology.

Confirming that increases in excitement and frustration during game speed increase events are related to flow would require additional examination and analysis that is beyond the scope of this project. However, the exploration of changes in affective responses to a single event over time as difficulty increases could prove a valuable avenue for future research.

### **5.3.2.2. *Non-Useful Tetriminos***

Players received a replacement Tetrimino immediately following the placement of their previous one within the game board. However, players had no control over the shape of the next Tetrimino they received, and often received a piece that was not suitable for the current state of their game board. Receipt of non-useful Tetriminos can hinder the ability of players to correctly place them upon the board, potentially inhibiting performance or resulting in mistakes being made. Likewise, placing Tetriminos either correctly or incorrectly has consequences for player choice and performance moving forward within the game.

As illustrated in Figure 90 below, arousal was minimally influenced by these events, but valence was reduced into a negative state, being one of only three game events within this study that elicited a mean negatively valenced affective response. Receiving non-useful Tetriminos had two main effects when compared against mean game play, namely decreased Fun, and increased Frustration (see Chapter 4: Results for full breakdown of results).

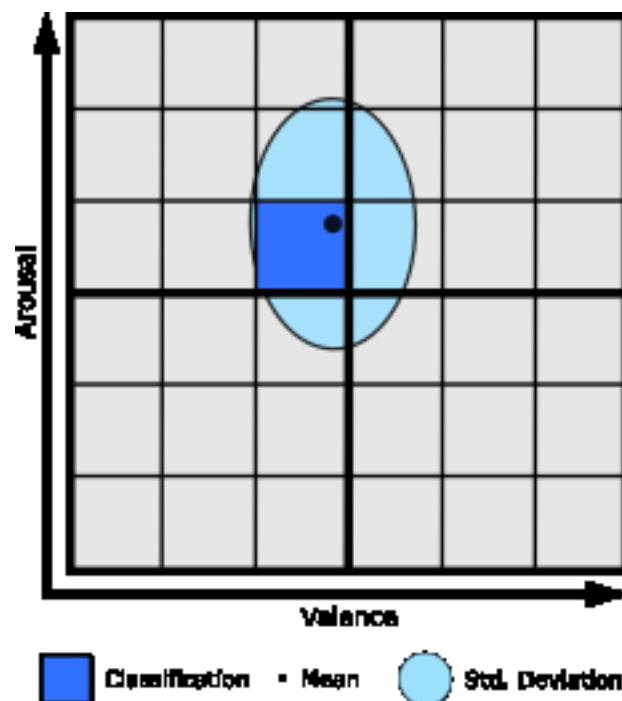


Figure 90: Affect Grid illustrating the mean Arousal and Valence for Non-Useful Tetriminos events.

Fun was found to decrease by an average of 7.99% from the mean game play experience during Non-Useful Tetriminos events, constituting a medium strength statistically significant effect. Simultaneously, Frustration was found to increase by 16.9%, which constituted a large strength statistically significant effect. As the measure for fun is modelled through having a high valence, and the measure of frustration modelled through having a low valence, it is logical that one should decrease while the other increases.

However, this does not necessarily imply that increases in difficulty should generally correspond with increases in frustration. Numerous studies have evidenced that increases in game difficulty and challenge are often associated with increased fun or enjoyment (Klimmt *et al.*, 2009; van den Hoogen *et al.*, 2012; Kneer, Elson and Knapp, 2016). Additionally, from the discussion of flow theory provided for game speed increase events (Section 5.3.2.1), it is expected that increases in difficulty should logically follow the theory and increase engagement with the game up until the point where the respective skills of the players are exceeded (Chen, 2007; Nacke and Lindley, 2009).

Upon initial examination, the results for non-useful Tetriminos do not appear to follow the results expected if flow theory holds true, with no associated mean increase in positive affect observed during these game events, nor any increase in modelled challenge. The most likely explanation is that near the start of the game, where game speed is lower and making mistakes is less damaging to overall performance, receiving non-useful Tetriminos presents little difficulty for the player to overcome, and therefore elicits no substantial influence on measures of Challenge from the mean. However, once game speed begins to increase and making mistakes inhibits performance, the player's skills will become stressed, after which point receiving non-useful Tetriminos may become more likely to elicit increased frustration or anxiety.

Within the current analysis of results, where all events of the same type are considered jointly rather than at varying points in time or under varying performance conditions, this appears as a predominately negative event. However, when event occurrences are examined individually as illustrated in Figure 91 below, it becomes evident that these events initially elicit minimal affective response, but subsequently become increasingly negatively valenced.

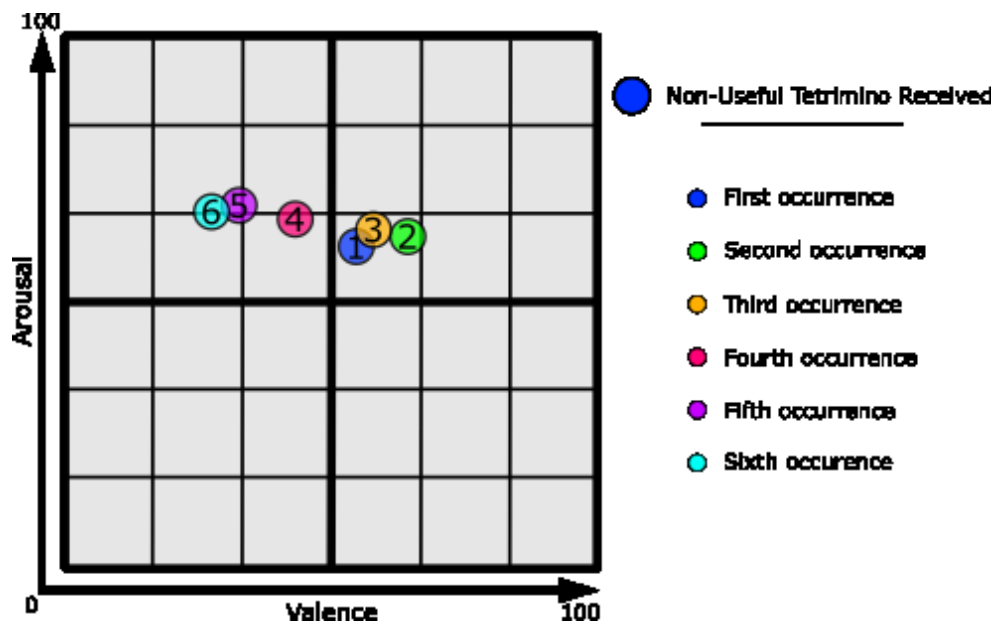


Figure 91: Non-Useful Tetriminos Received events placed in order of occurrence for a single participant, illustrating changes in event-response over time.

This is congruent with flow theory, as under low stress it is likely that players would have the time and cognitive resources available to identify locations on the board to deal with non-useful Tetriminos without too much difficulty. However, as the game speed increases, players have less time available to make decisions, straining cognitive resources (Nacke and Lindley, 2009; Jin, 2012; Klasen *et al.*, 2012). It is likely that this cognitive strain is experienced as a frustration rather than as a challenge, especially if players are not experienced at playing Tetris, as performance is largely based upon making decisions very quickly. Aligning with flow theory, players without the skills necessary to make these rapid decisions and movements would quickly be overwhelmed (Jin, 2012; Liu and Chang, 2012).

Additionally, this aligns with the theory of cognitive workload, whereby performance of cognitively engaging tasks result in fatigue and eventually frustration once the cognitive capacity of the individual has been met for extended periods of time, or exceeded (Koeffel *et al.*, 2010; Cowley *et al.*, 2016).

### 5.3.2.3. Correct Placement (Player + Opponent)

Placing Tetriminos in locations that were appropriate for their respective shapes, and did not block any area of the game board were coded as correctly placed, as they did not inhibit performance moving forward within the game. These events were recorded both when the player themselves correctly placed Tetriminos, and also when the player observed their opponent correctly placing Tetriminos. This resulted in subtly different affective responses to both classifications of events despite them being the same type (see Figure 92 below for a comparison).

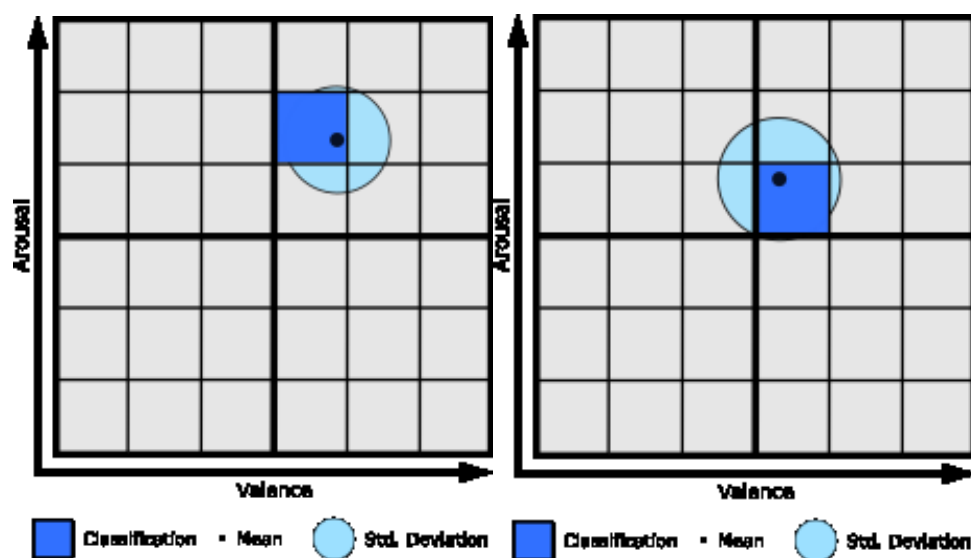


Figure 92: Affect Grid modelling the mean arousal and valence scores occurring during Correct Tetriminos Placement events when completed by Players (Left) and Opponents (Right), illustrating that the affective responses were moderately arousing and positively valenced.

In the case of players correctly placing Tetriminos, all five modelled emotions were influenced to varying degrees. Both Fun and Excitement were found to increase when Tetriminos were correctly placed (8.49% and 11.44% respectively), while Boredom, Frustration, and Challenge were found to decrease (-8.77%, -4.54%, and -7.39% respectively) (see Section 4.2.2.2.2 for a full breakdown of results).

Correctly placing Tetriminos represented a positive feedback mechanic within Tetris Party Deluxe, as players could see visually whether their respective placements were

appropriate or not, depending on the current state of their game board. Feedback mechanics are a common feature of video games, as they allow players to assess their own performance (Klimmt *et al.*, 2009). The results of events related to correctly placing Tetriminos therefore fit within two dominant theories explaining human motivation, namely Self-Determination Theory (Ryan, Rigby, and Przyblyski, 2006), and Attribution Theory (Weiner, 1985).

The Self-Determination Theory (SDT) of motivation is perhaps the most widely accepted and validated theory of motivation within the video game research domain (Hamari, Koivisto and Sarsa, 2013; Deterding, 2012; Deterding, Nixon, Khaled, and Nacke, 2011). Most important for this result is that humans are said to be innately driven by a need to feel competence, to overcome and master challenged (Deci and Ryan, 2008). Feedback mechanics have previously been found to directly influence feelings of competence and mastery (Vansteenkiste and Deci, 2003; Potts *et al.*, 2011; Lyons, 2015), which have likewise been linked with increased game enjoyment, including joy and fun (Vorderer, Klimmt, and Ritterfeld, 2004; Klimmt *et al.*, 2009).

SDT would suggest that emotions related to fun and joy would increase following events contributing towards competence and mastery, which the results for this study support. Additionally, it is likely that increased perceptions of competence and mastery would coincide with decreased emotions related to frustration and challenge, as it is likely that players would perceive themselves as being more capable following receipt of positive feedback (Przyblyski *et al.*, 2014; Crutzen, van 't Riet and Short, 2015). The affective responses found when players correctly place Tetriminos therefore appear to be congruent with feedback events influencing feelings of competence and mastery.

In the case of players observing their opponents correctly placing Tetriminos, only Challenge differed significantly from the mean game play, with a decrease of -7.39% (see Section 4.2.2.4.1 for a full breakdown of results). This is interesting primarily



because observing other players succeeding could be expected to increase the perceived difficulty for the player if they wish to win the game, which may be expected to increase modelled emotions relating to Challenge. It could also be expected that increasing the competitive difficulty in this way may increase Frustration, if beating the opponent was perceived as beyond the skills of the player.

It appears likely that the lack of modelled changes may be related to the lack of direct competitive interactions between the player and their opponents, being that they are not able to directly influence the performance of the other player, which is a core aspect of typical competitive play (Vorderer, Hartmann and Klimmt, 2003; Tauer and Harackiewicz, 2004; Marker and Staiano, 2015). This would mean that despite opponents undertaking an action that may place them at a competitive advantage, the lack of actual interaction may have caused players to largely ignore these events. If this holds true, then the decrease in modelled Challenge may be related to an external factor rather than to the experience of these events. This would be an interesting avenue for future research to be conducted into the influence that player interactions within competitive contexts has upon challenge.

### **5.3.2.4. *Incorrect Placement (Player + Opponent)***

Placing Tetriminos in locations that were not appropriate for their respective shapes, and where they blocked areas of the game board or otherwise inhibited player performance moving forward in the game were coded as being incorrectly placed. These events were recorded both when the player themselves incorrectly placed Tetriminos, and when the player observed their opponent incorrectly placing Tetriminos. This resulted in subtly different affective responses to both classifications of these events, despite them being of the same type (see Figure 93 below for a comparison).

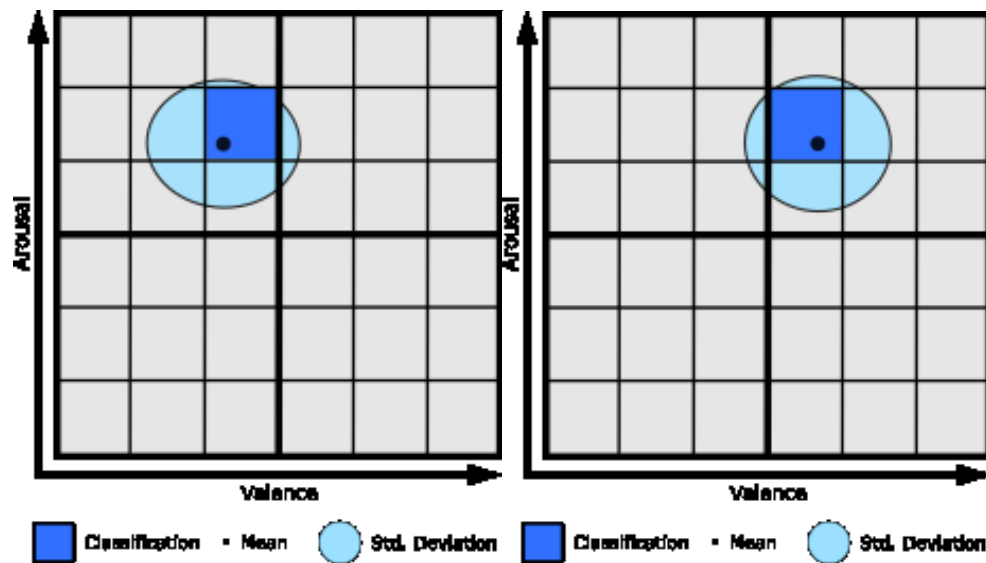


Figure 93: Affect Grid modelling the mean arousal and valence scores occurring during Incorrect Tetriminos Placement events when completed by Players (Left) and Opponents (Right).

In the case of players incorrectly placing Tetriminos, four modelled emotions were influenced to varying degrees. Excitement, Frustration, and Challenge all increased (11.78%, 27.12%, and 17.81% respectively), while Boredom decreased by -5.04% (see Section 4.2.2.2.3 for a full breakdown of results). As discussed in the results for correct placement of Tetriminos (Section 5.3.2.1 above), the placement of Tetriminos serves as a feedback mechanism, as players can visually assess their respective placements and performances depending upon the state of their game board.

The strength of these differences is likely related to incorrect placements always being experienced as negative events irrespective of the current game speed or difficulty. This contrasts against other events influencing difficulty, such as non-useful Tetriminos being received, which appear to elicit varying affective responses depending upon the relationship between the ability of players and the current difficulty of the game.

The differences in affective responses elicited between these types of negative events may be explained through two phenomena. Firstly, placing Tetriminos incorrectly was frequently observed to be followed by expressions of exasperation or disappointment. This suggests that making mistakes may have caused players to become frustrated with their respective performance. Secondly, irrespective of player frustration due to making errors, Tetriminos being placed in poor locations can block off gaps in the game board, making it more difficult to clear lines and thereby inhibiting performance. These results fit both with the Self-Determination Theory (SDT) as it relates to player motivation (Ryan, Rigby, and Przyblyski, 2006), and the flow theory as it relates to video game engagement (Nacke and Lindlay, 2009).

Players becoming frustrated with their performances when making mistakes fits SDT due to providing negative feedback that undermines feelings of competence, and consequently negatively influencing enjoyment of the game experience (Przyblyski, Rigby, and Ryan, 2010). This suggests that players would likely experience these events through negative affective responses irrespective of their outward expressions of disappointment.

As previously discussed, flow theory posits that the optimum of fun and engagement occurs when an individual's skills are matched with the difficulty of the task. However, placing Tetriminos incorrectly increases the difficulty of the game, which consequently increases the likelihood of misplacing Tetriminos in the future. Depending on the current state of the game board and game speed, misplacing Tetriminos could therefore result in a rapid escalation of game difficulty to a level beyond the skill level of players, consequently undermining self-esteem, leading to frustration and sadness (Klimmt *et al.*, 2009).

In the case of players observing their opponents incorrectly placing Tetriminos, Fun increased with a relatively small but statistically significant effect of 5.97%, while Boredom, Frustration, and Challenge all decreased with relatively large effects (-

8.7%, -12.13%, and 18.47% respectively) (see Section 4.2.2.4.2 for a full breakdown of results). This is interesting primarily because it does not appear to follow Attribution Theory, which posits that enjoyment of competitive tasks stems from the ability of individuals to attribute success to their own performance and skills, rather than from the misfortune of others (Klimmt *et al.*, 2009).

This also directly contrasts against the affective responses that players had to their opponents correctly placing Tetriminos, where they had no influence upon emotions relating to enjoyment, and only a small influence upon challenge. One possible explanation could be that players paid more attention to the failures of their opponents, but minimal attention to their successes. As previously stated, players were observed to regularly express exasperation or disappointment when they incorrectly placed Tetriminos, which likely draws the attention of the other player to the event.

This ties in with the concept of *schadenfreude*, defined as finding pleasure in the misfortune of others, and often manifests in games as taking pride in beating an opponent, or gloating over misfortunes that occur to them (Schell, 2014). It is plausible therefore that the increased positive emotions that players experience when their opponents misplace Tetriminos is not due to it improving their relative competitive advantages, but rather due to experiencing *schadenfreude*.

### 5.3.2.5. Multiple Line Clear (Player + Opponent)

Clearing multiple lines of Tetriminos is a high-performance event where players successfully cleared between two to four lines of Tetriminos from the game board by placing a single Tetrimino, usually requiring forethought and planning to accomplish. These events were recorded both when the player themselves cleared multiple lines of Tetriminos, and when the player observed their opponent clearing multiple lines. This resulted in different affective responses to both classifications of these events, despite them being of the same type (see Figure 94 below for a comparison).

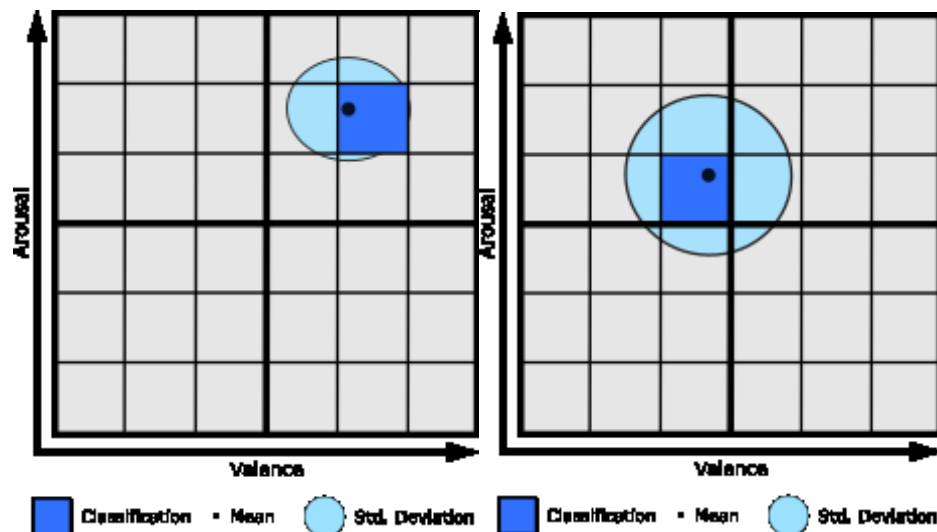


Figure 94: Affect Grid modelling the mean arousal and valence scores occurring during Multiple Line Clear events as experienced by Players (Left) and Opponents of Players (Right).

In the case of players clearing multiple lines of Tetriminos, all five modelled emotions were influenced to varying degrees. Fun and Excitement both increased significantly (9.27% and 9.93% respectively), while Boredom, Frustration, and Challenge all decreased significantly (-13.81%, -9.86%, and -16.12% respectively) (see Section 4.2.2.3.1 for a full breakdown of results). These results reflect the results for correct Tetrimino placement, albeit with larger effects against mean game play. Additionally, players successfully clearing multiple lines of Tetriminos were regularly observed to make expressions of joy or excitement.

As with the correct placement of Tetriminos, clearing multiple lines of Tetriminos constitutes a positive feedback mechanism through which the player can monitor their own performance, which in this event is positive. While multiple line clear events elicit stronger positive affective responses due to being an indicator of higher performance, the mechanisms through which these events work appear to be the same as discussed for the correct placement of Tetrimino events, with multiple line clear events likewise encouraging feelings of competence and mastery, as postulated by SDT.

In the case of players observing their opponents incorrectly placing Tetriminos, Frustration increased by an average of 6.36% from the mean, and Boredom decreased by an average of -9.41% from the mean (see Section 4.2.2.4.2 for a full breakdown of results). These contrasts unambiguously against Correct Placement of Tetriminos events, where players observing their opponents successfully placing Tetriminos experienced only a slight decrease in emotions relating to Challenge. As with Incorrect Placement events, a key reason for this might be that players successfully clearing multiple lines of Tetriminos were likely to make expressions in reaction to the events, ranging from minor vocalisations to asking if the other player saw what happened.

This draws the attention of the player to the event, increasing the likelihood that the player will compare their relative performances and adding to the perceived difficulty for the player if they wish to win the game. Perhaps most interesting here, especially in comparison against events where the opponents correctly placed Tetriminos, is that Frustration increased while Boredom decreased, but again there was no increase in Challenge that might be expected in competitive games. One possible explanation could be that these events may have been experienced as annoyances or distractions, rather than as increases in difficulty.

Comparing against events where opponents correctly placed Tetriminos, for which players were observed glancing at the opponent as they completed these events but did so of their own volition, the difference may well be elicited by the opponent speaking or expressing positive emotions. It is possible that this expression may draw the attention away from the player at a critical moment, or otherwise break their concentration or flow state. If this is the case, then the negative affective responses associated with these events may be due to the event being regarded as a distraction rather than as a game experience, which has been found to negatively impact game experiences and enjoyment (Smith and Pittman, 1978; Jegers, 2007; Ceaparu *et al.*, 2004). This would also explain why Challenge did not increase, as the

player may not respond in such a way that would reflect increased emotions relating to difficulty.

### 5.3.2.6. Player Win vs Player Loss

As a competitive video game, the objective of Tetris Party Deluxe is to overcome the opponent player. The player is therefore able to either Win or Lose the game, depending upon their performance relative to their opponent. Affective responses to the player winning versus the player losing is an interesting consideration, and so both series of events were recorded and differentiated, with considerable differences observed (see Figure 95 below for a comparison).

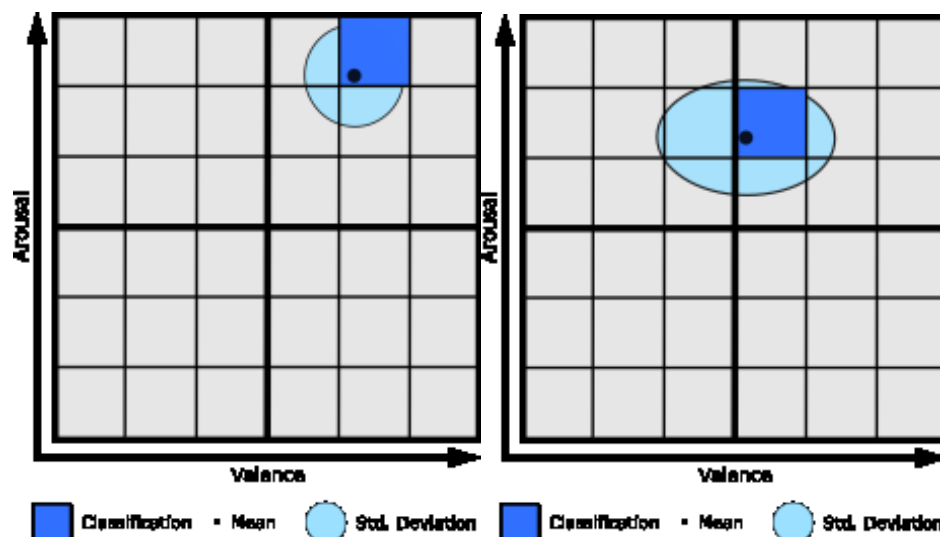


Figure 95: Affect Grid modelling the mean arousal and valence scores occurring during Player Win (Left) and Player Loss (Right) events, illustrating that the affective responses were considerably different.

In the case of players winning the game, four of the modelled emotions were influenced significantly. Measures of Fun and Excitement increased by 8.83% and 24.17% respectively, while Boredom and Frustration were reduced by -12.09% and -10.59% respectively (see Section 4.2.2.3.2 for a full breakdown of results). These results are not at all surprising, given that the player has just achieved the goal of the game, and beaten their opponent.

Video games are generally goal oriented, which holds especially true for competitive games where the goal is to overcome opponents (Andrew K Przybylski, Rigby and Ryan, 2010). A wide body of work highlight the positive emotional experiences associated with general winning experiences (Vansteenkiste and Deci, 2003; Tauer and Harackiewicz, 2004; Marker and Staiano, 2015), overcoming challenges (van den Hoogen *et al.*, 2012), or overcoming opponents in competitive interactions (Tauer and Harackiewicz, 2004; Marker and Staiano, 2015), and the results of winning events within this study appear to be congruent with the position that winning games typically elicits strong positive emotions such as joy and excitement.

In the case of players losing the game, four of the modelled emotions were influenced significantly. Measured Frustration were increased by 17.08% from the mean, while measures of Fun, Boredom, and Challenge decreased by -18.31%, -4.34%, and -6.23% respectively. While decreases in Boredom and Challenge were significant, these represented relatively small changes overall. Most interesting is the relatively large changes in Fun and Frustration, which also had larger than normal variances amongst participants.

Additionally, while no win events were experienced as negative affective responses, most loss events were also experienced as positive affective responses. As illustrated in Figure 95 above, the standard deviation for valence responses under Player Loss events is much larger than that for the Player Win events. These variations may be explained through two phenomena likely related to the relationship between participants, namely the nature and quality of the friendship between player dyads, and the relative difference in skill between players.

For the first, although the nature and quality of friendships was neither controlled for (participants were recruited in pairs, but no qualifier was attached) nor analysed against, it appears likely that the nature of the relationships could elicit greatly varied responses to the game outcome. For example, it was observed that some



participants experienced positive affective responses when their opponent performed well, including expressing joy and other associated positive affective responses when their opponent won the game. As this was not controlled for nor directly analysed within this project, and as existing studies examining social video game play typically assume equal friendship quality and do not attempt to control for types of relationships (e.g. Mandryk and Atkins, 2007; Peng and Hsieh, 2012; Marker and Staiano, 2015), this could be an interesting avenue for future research.

The second phenomena that could explain the large variations in valence of affective responses is the relative closeness of the game. In game sessions where both participants were closely matched in skill, game loss events appeared to elicit far less frustration than those where skill levels were mismatched, which elicited far more frustration generally. This is generally consistent with existing studies that demonstrate that enjoyment from video game play does not necessarily stem from winning or losing, but rather from challenging game play (Nacke and Lindlay, 2009; Ravaja *et al.*, 2005).

An additional factor that may influence the positive affect response to losing the game comes from literature on gambling theory, namely on the psychology of near miss losses (Banks *et al.*, 2017). Gambling theory illustrates that individuals experience positive emotions even when they lose a bet, if they perceive that the outcome was nearly in their favour and that they only narrowly missed out on winning (Banks *et al.*, 2017; Dixon *et al.*, 2011). There is evidence of this theory at work in video game studies, where players often report higher levels of satisfaction and enjoyment when the game is perceived as fair and within their grasp (Kong *et al.*, 2011; Rigby & Ryan, 2011). However, these studies provide data generated from self-report instruments after several minutes have passed, potentially providing time for players to reflect upon their performance and on the game generally. To date, no studies to date appear to have examined the phenomena of the immediate affective response to a game loss being positive, which appears relatively counter-intuitive as an automatic response to loss.

Irrespective of winning or losing, for all game sessions observed within the study, participants immediately began to talk animatedly and discuss the game outcome, typically with friendly banter exchanged between participants. This highlights that even negatively valenced affective responses, such as frustration at losing, may be typically experienced within the context of video game play generally, meaning playing video games with friends is generally a positive experience, regardless of whether the player lost the game at the end. This result certainly appears to support the often-made argument for the magic-circle of video game play, where players exist within a bubble separated from real-world experiences, and do not always respond in the ways that they might otherwise.

### 5.3.3. Summary of Research Question 1

Research Question 1 aimed to explore the influence that identified video game events can have upon affective responses as experienced during real-time game play. The unit of measurement was therefore observed game events, analysing the influence that these had upon affective responses during game play. For each game event, a discussion was provided regarding the outcomes of five modelled emotional responses, namely Fun, Excitement, Boredom, Frustration, and Challenge (see Section 3.8 for further details on the modelling process).

Affective responses were discussed in relation to existing psychological theories as they relate to game experiences, such as the Flow Theory of engagement, and the Self-Determination Theory of motivation, and ties in previous studies conducted within the domain of video game research. The affective responses reported within this study appear to be largely congruent with previous studies, such as events that increase difficulty largely resulting in increased frustration.

Additionally, some avenues for future research have been identified, such as the influence that social interactions that distract the player can have upon frustration,

or the moderating influence of relative performance of players at the time of events occurring.

Overall, the volume and relative quality of real-time data generated on the affective responses to observed events occurring during video game play supports the usefulness of the method used within this study. As this study was conducted as an exploration of game events and the affective responses that may occur in response to them, the scope of this study limited the ability to draw specific in-depth insights about each result and for each event identified. However, studies aiming to develop insights into more specific game events or targeted affective responses ought to be able to use this method to examine events in far greater detail.

### 5.4. Research Question 2

#### ***Research Question 2:***

*How does the use of psychophysiological instruments influence the player experience during video game play?*

Research Question 2 aimed to explore the influence that the use of psychophysiological equipment, when placed upon the body of participants, had upon the game experience, both in terms of general perceived game experience, and affective experience. The unit of analysis was therefore player experience, analysing the influence that physiological measurement devices had upon self-reported affect and game experience.

This required the collection of subjective data pertaining to relevant game experience factors, namely the Game Experience Questionnaire (GEQ) to measure general perceived game experience, and the Self-Assessment Manikin (SAM) to measure affective experience. Additionally, videos of participants were used as a secondary measurement of facial expressions, allowing for an examination of

whether the use of psychophysiological equipment influenced the elicitation of facial expressions.

Previous studies utilising psychophysiology within video game contexts have demonstrated that physiological data is able to provide valid and reliable insights into game experiences and mean emotional experiences of game play (e.g. Ravaja and Kivikangas, 2008b; Drachen, L. E. Nacke, *et al.*, 2010; Kivikangas *et al.*, 2011; Nogueira, Rodrigues and Oliveira, 2013; Cowley *et al.*, 2014; Klarkowski *et al.*, 2016). The comparisons between subjective and objective data in this study are therefore conducted with the intention of examining consistency of this method against existing approaches, rather than to prove validity.

The primary intention of this discussion is to examine the influence that presence and use of psychophysiological equipment was found to have had upon game experiences. Despite the apparent correlation between objective and subjective measures used within this study, there were a series of measured effects that could have a range of implications for future studies utilising psychophysiology, and these are discussed in greater detail within the following sections.

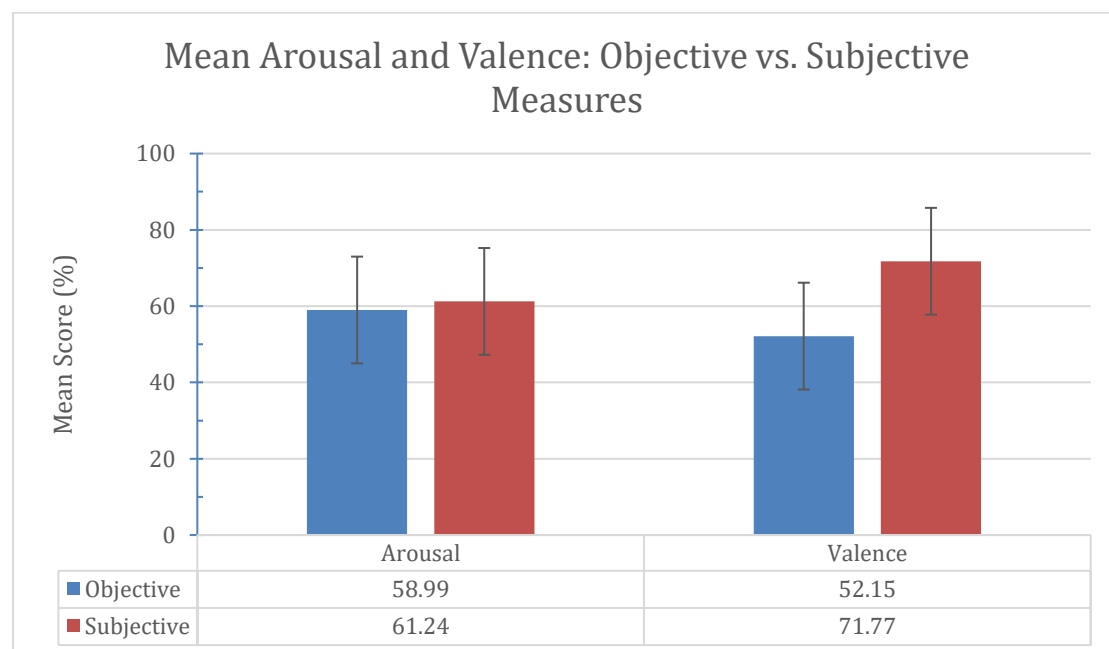
### 5.4.1. Relationship between Subjective and Objective Measures

Subjective survey instruments are by far the most common tool used for the assessment of video game experiences, largely due to being relatively simple to deploy and interpret. Despite their ubiquity within the domain of video game research, survey instruments have a series of limitations that inhibit the collection of data surrounding some phenomena. These limitations, which are discussed in greater detail in Chapter 2, include the inability to collect nuanced data on the experiences that occur during game play, as they are typically conducted at the completion of a game experience.

The use of objective psychophysiological measures provides an avenue for collecting nuanced data as it occurs, and so facilitates the examination of event-related data

(Cacioppo, Tassinary and Berntson, 2007; Cowley *et al.*, 2016). However, as the objective data can be relatively ambiguous and relies upon a substantial degree of processing and interpretation, the assessment of objective data should be correlated against valid subjective instruments (Cacioppo and Tassinary, 1990; Cacioppo, Tassinary and Berntson, 2007; Cowley *et al.*, 2016).

To this end, as discussed in Section 3.5.2.2, the Self-Assessment Manikin (SAM) survey instrument (Morris, 1995) was therefore selected and used to collect data on the perceived arousal and valence of participants at the completion of game play. As evidenced in Figure 96 below, the two measures of arousal are highly similar with a difference of only 2.25, while the two measures of valence differ by 19.62 (see Section 4.3 for a full breakdown of results). For comparison purposes, the subjective SAM scores were converted to percentage of their maximum, as this is the format that the objective scores were in.



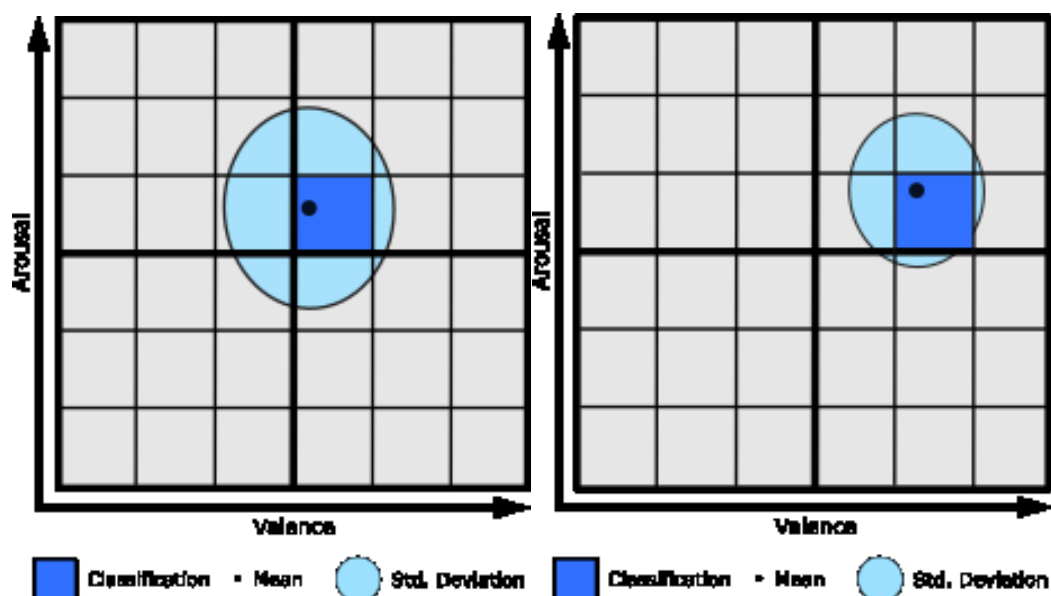
*Figure 96: Chart illustrating mean differences in Arousal and Valence for both Objective and Subjective measures.*

A common problem with subjective surveys is scaling of questions, where participants will score items based on their current frame of reference rather than

on absolute scores, which has presented issues in previous studies attempting to correlate objective and subjective scores (e.g. Mandryk *et al.*, 2007). To use an example provided by Mandryk and co-authors (2007), asking participants to rate their frustration for a game experience may result in a relatively high value within the context of the game, as the game experience is the frame of reference. However, the objective measures are reported as percentages of maximum possible values, making comparisons and correlations difficult.

The SAM survey attempts to control for this by presenting images of humanoid figures on a scale between absolutes, such as the figures for Arousal, which are depicted on one extreme with a heart that is exploding within the chest, and on the other extreme as having an extremely small heart (see Section 3.5.2.2).

A Pearson's test of correlation revealed that the two measures of Arousal were highly correlated, with a statistically significant result indicating that the subjective measure of Arousal explains approximately 80% of the objective measure of Arousal. A second Pearson's test revealed that while the correlation was more moderate, the two measures of Valence still feature a statistically significant correlation, with the subjective measure of Valence explaining approximately 25% of the objective measure of Valence.



*Figure 97: Affect Grids illustrating differences between the mean Arousal and Valence scores for both Objective Measures (Left grid) and Subjective Measures (Right grid).*

As illustrated in Figure 97 above, despite high levels of correlation between subjective and objective measures, when modelled onto the AV space grids there are still substantial differences evident. While these issues may be due to scaling issues between the subjective and objective measures, which would be consistent with other studies attempting to correlate objective and subjective measures of emotion (e.g. Mandryk, Atkins and Inkpen, 2006; Drachen, L. E. Nacke, *et al.*, 2010), the potential for these differences to be related to the use of physiological equipment is discussed in Section 5.4.2 of this discussion chapter.

As an additional note, the standard deviations also differ substantially between the two measures. The objective measures appear to elicit a much larger variation of scores amongst the participant population, while subjective scores appear to be more homogenous. This is likely because of a lack of granularity within subjective surveys, as the SAM survey offers only nine levels of granularity, compared to the thousands of levels available through the objective measures.

### 5.4.2. Measurement Effects

Measurement effects are a ubiquitous aspect of research across domains, whereby the act of observing or measuring phenomena often influences the properties observed (Kassam and Mendes, 2013). In the case of research into video game experiences and emotions, both subjective and objective instruments of measurement are likely to exert an influence upon the experiences of game play being examined.

Emotions are inherently complex, involving multiple physiological systems that interrelate and exert varying influences upon each other depending upon the emotional state and situational contexts (Scherer, 2005; Kassam and Mendes, 2013).

A core aspect of emotional experiences, particularly as they relate to video games, is that while individuals may be passively aware of emotions, the act of appraising or reflecting upon the emotional experience disrupts these processes (Lambie and Marcel, 2002; Kassam and Mendes, 2013). Additionally, emotional experiences are highly subjective in nature, and require the individual to interpret them, which may differ substantially between individuals (Robinson and Clore, 2002; Kassam and Mendes, 2013).

There is a wide body of evidence that reflecting upon emotional experiences has a substantial impact upon the current emotional states of individuals (Frattaroli, 2006; Gross, 2002), and the ability to do so serves as the basis for many psychological therapy techniques (Kassam and Mendes, 2013). These influences occur even with relatively simple reflections upon emotions, such as simply being made aware that images will be about positive or negative emotions causing a change in participant ratings of images illustrating varying emotions (Jakymyn and Harris, 2012), or being asked to rate an image as either happy or sad influencing subsequent neural activities (Lieberman *et al.*, 2007).

The most common instruments used to measure emotions and experiences within the domain of video game research are surveys requiring reflection upon emotions or experiences. However, given the above discussion it is expected that the use of these instruments exert a substantial influence upon the reported outcomes. These influences of subjective instruments form a significant portion of the value behind psychophysiological assessment of experiences and emotions, as psychophysiology does not require the participant to be conscious of or reflect upon their emotions (Cacioppo, Tassinary and Berntson, 2007; Drachen, L. Nacke, *et al.*, 2010; Kivikangas *et al.*, 2011).

However, while the use of psychophysiology and associated instruments for the measurement may be discrete and covert in terms of recording data relating to the emotions and experiences of individuals (Cacioppo, Tassinary and Berntson, 2007;



Ravaja and Kivikangas, 2008b), the requirement to place electrodes upon the body of individuals is likely to also influence the experience of participants. Within the context of research into video games, this likely further compounds existing differences between natural video game play and video game play as conducted within controlled experimental settings.

Experimental settings typically require participants to play a specific game (rather than opting to play of their own volition) within an environment that is different to where they might normally play (e.g. in the comfort of their home), and where they are aware of measurements or observations being taken (rather than having privacy).

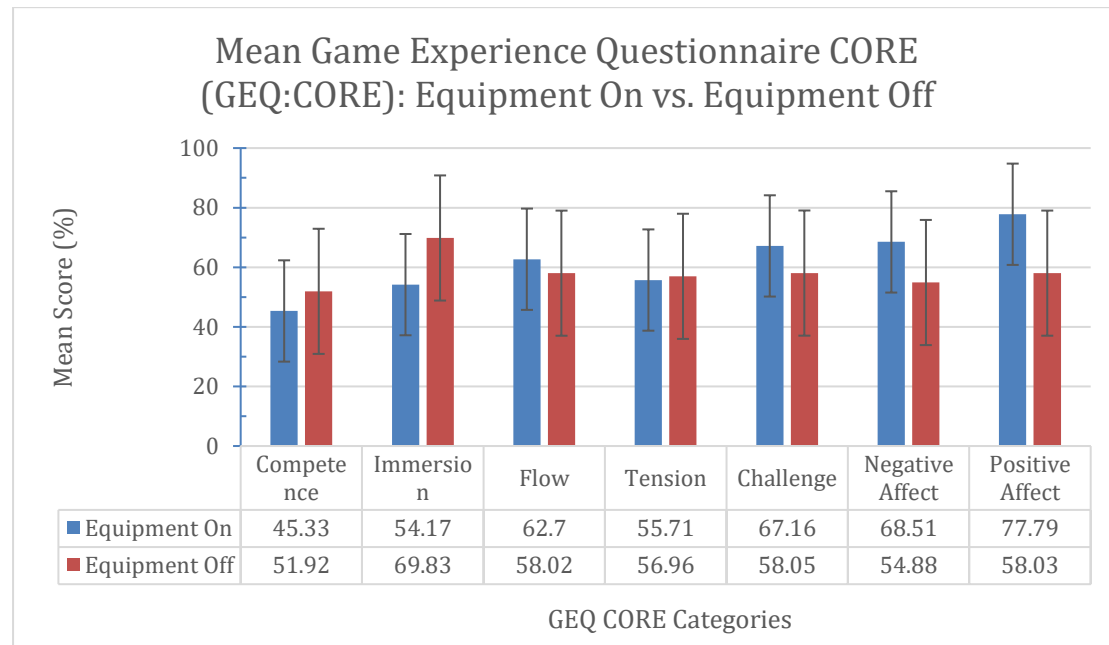
Within the domain of video game research, some studies using psychophysiology attempt to control against the influence that the equipment may have upon participants, such as Mandryk *et al.* (2007) who attached equipment to both individuals in social game play but only recorded measures from one of the two individuals, partly in an attempt to mitigate the influence that the equipment could have. However, the clear majority of studies ignore the potential measurement effects that the equipment could have upon the game experience, potentially due to the specific measurement effects not being known, or perhaps because they assume that measurement effects must exist but are unavoidable.

This provides the basis for the value behind Research Question 2 of this project, where the measurement effects of psychophysiology within a video game context are explored in terms of the influence upon perceived game experiences, affective experiences, and facial expressions.

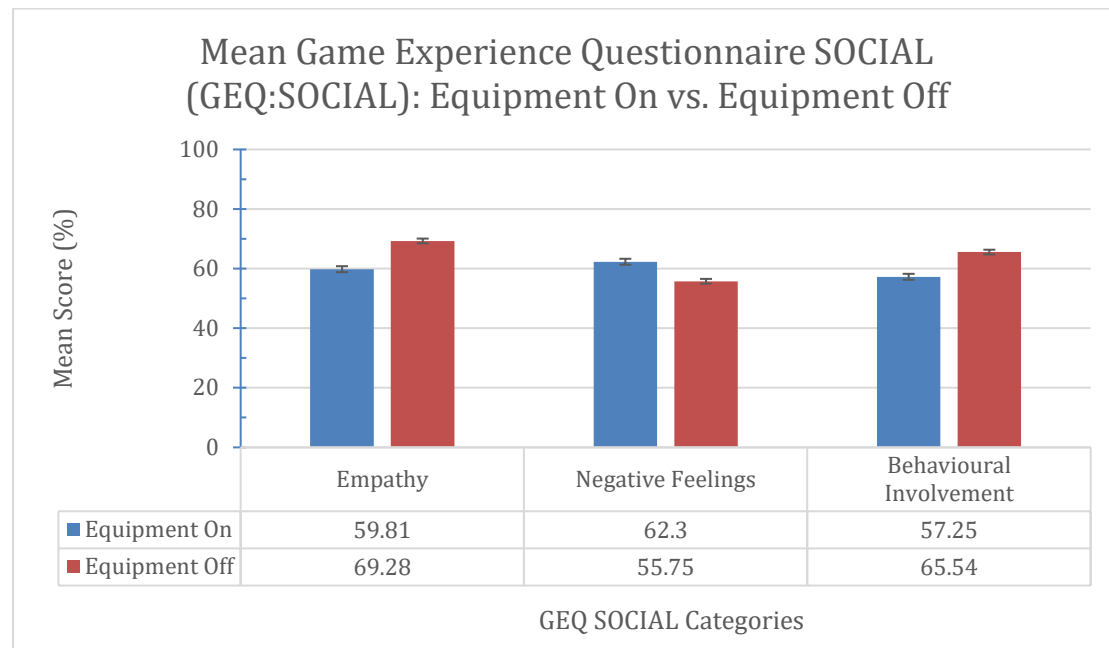
### **5.4.2.1. Measurement Effects upon Game Experience**

The Game Experience Questionnaire (GEQ) was used to collect data on the perceived game experience of participants, providing insights into a range of experiences that relate to video game play. As this study used a social method, both the core

experience and social modules of the GEQ were deployed, facilitating additional exploration of a broad range of potential influences. As outlined within the study procedures (Section 3.7), GEQ results were split between Equipment On and Equipment Off experimental conditions, with several differences in experience becoming evident (see Figure 98 and Figure 99 below).



*Figure 98: Chart illustrating mean differences between Equipment On and Equipment Off conditions for each of the GEQ CORE components.*



*Figure 99: Chart illustrating mean differences between Equipment On and Equipment Off conditions for each of the GEQ SOCIAL components.*

As presented in Section 5.4.1, the presence of psychophysiological equipment resulted in four significant effects upon the core game experience as reported by the CORE GEQ module. Self-reported Immersion was found to decrease by an average of -15.66% when equipment was used, while perceived Challenge, Negative Affect, and Positive affect were all found to increase (by 9.12%, 13.63%, and 19.76% respectively). These results appear to indicate that the use of equipment disrupted the game play experience, serving to reduce the immersion of participants during game play, and increasing their apparent awareness of their emotions.

Additionally, as presented in Section 4.3.2, the presence of psychophysiological equipment resulted in two significant effects upon the social game experience as reported by the SOCIAL GEQ module. Feelings of Empathy towards the other player were found to reduce by an average of -9.48% when psychophysiological equipment is used, while perceptions of Behavioural Involvement, which are related to the perceived ability to interact socially within the game, were reduced by an average of -8.29%.

The decrease in Immersion and increase in Challenge are therefore likely explained through the diversion of participant attention away from game play itself. Player attention is an essential component of performance, with games requiring a high degree of attention to perform effectively. Decreased Immersion may therefore be indicative of decreased player attention to the game, which could negatively influence player performance. This decreased performance may subsequently increase the perceived Challenge of the game, since it would be harder for players to successfully undertake actions within the game when distracted.

When considered alongside the increases in perceived Negative and Positive Affect, this may indicate that participants instead reflected upon their emotions at the expense of their respective immersion. The apparent reduction in perceptions of Empathy and Behavioural Involvement relevant to the social aspects of game play add weight to this argument, as this could suggest that participants under the Equipment On conditions may have become introspective and less socially responsive.

The lack of significant differences in perceived Tension, Flow, and Tension may suggest however that perceived performance may not have been impacted by these distractions or introspective states, though future research may prove interesting to assess whether actual performance was influenced regardless of perceptions.

These results are similar to results reported by Kassam and Mendes (2013), where they indicate that participants being asked to reflect upon their emotions served to increase perception and experience of those emotions. This could suggest that, despite participants not being explicitly told that the psychophysiological equipment would be recording data pertinent to emotions, participants nonetheless became more aware of their emotions, and reflected more upon them.

As the GEQ was administered at the conclusion of game play sessions, an alternative possibility is that participants did not differ during game play, and that these

reflections were due to participants becoming aware of the psychophysiological equipment when undertaking the surveys, resulting in the differences observed. If this holds true, then it could explain the increases in Negative and Positive Affect as this is congruent with existing work demonstrating the influences of reflecting upon emotions (Kassam and Mendes, 2013). However, it does not necessarily hold that participants would retrospectively perceive the game to be less immersive due to these reflections, as increased emotional experiences are more typically associated with increased perceived immersion (Ermi and Mäyrä, 2005; Nacke and Lindley, 2009; Andrew K. Przybylski, Rigby and Ryan, 2010).

A limitation on the strength findings presented here for RQ2 is the use of the GEQ instrument, which recent studies have highlighted as potentially problematic in terms of its validity for some of the items it measures (Brühlmann & Schmid, 2015; Abeele et al., 2016). Brühlmann and Schmid (2015) found that some items measuring the factors for Challenge and Tension in the GEQ were not strongly associated with the game experiences that might be expected from these items within the scale. Therefore, while the GEQ results in this project clearly suggest that there were influences upon the game experience due to the presence of the equipment, the results cannot be interpreted with a strong degree of certainty. Further research should be undertaken with a range of additional measures, including open-ended interviews, to further develop an understanding of the specific influences that the presence of recording instruments has upon participants and their respective game experiences.

Undertaking a Factor Analysis of the subjective results was beyond the scope of this research project, as the intention was to provide a preliminary investigation into whether there is an influence and provide some early insights into where those may exist. The data generated within this study, while providing a valuable insight, is not sufficient for strong conclusions or even for the conductance of a Factor Analysis, as these typically require a larger participant sample to provide strong results. Further research should seek to examine the specific factors of research instruments that are influenced by the presence of physiological measures, and would require a larger participant sample, additional quantitative and qualitative instruments, and additional analytic methods to provide stronger and more holistic insights into these measurement effects.

Given the measured influences upon social interactions, facial expressions, and self-reported player experiences, it is therefore considered probable that the use of psychophysiological equipment did serve as a distraction during game play, diverting participant attention away from the video game experience and towards an introspective reflection upon their emotions. However, these results should be considered preliminary only, and further studies should be undertaken to confirm the influences in other environments and with additional instruments. It would therefore be interesting for future research to examine the influence of non-contact physiological measures, such as the measurement of heart rate through cameras recording participant faces (Muender *et al.*, 2016), and automatic facial pattern recognition from video feeds (Calvo and D’Mello, 2010; Sariyanidi, Gunes and Cavallaro, 2015). This would facilitate an examination of whether these influences are related to the use of electrode contact, or merely the knowledge that they are being observed.

### 5.4.3. Summary of Research Question 2

Research Question 2 aimed to explore the influence that the use of psychophysiological equipment, when used during experimental sessions of video game play, had upon the perceived game experience of participants. The unit of measurement was therefore player experience, analysing the influence that physiological measurement equipment had upon self-reported game experiences.

The discussion of RQ2 began with a comparison between subjective and objective measure of arousal and valence, facilitating an exploration of the relationship between the measures used within this study, demonstrating that both subjective and arousal measures were significantly correlated. These correlations align with previous bodies of work that combine objective and subjective measures, adding further weight to the argument that objective psychophysiological measures can be used effectively for the examination of affective responses to video game play.

The measurement effect of psychophysiological equipment was then discussed through the examination of subjective self-reports of game experiences, comparing differences observed between the Equipment On and Equipment Off experimental conditions, with several significant differences identified and discussed. These influences have a number of potential implications for future research, and warrant further examination to confirm these effects under varying game play conditions.

Overall, the subjective and objective measures used within this study indicated that psychophysiology was an effective tool for the examination of emotions and game experiences. As this study was conducted as an exploration of game events and the affective responses that may occur in response to them, the scope of this study limited the ability to draw specific in-depth insights into the explicit nature of measurement effects under varying conditions. However, the results of this study indicate that there do appear to be significant measurement effects that could impact upon the results and interpretation of psychophysiological results in future studies. As such, further examination of measurement effects could prove to be a highly valuable avenue for future research within the domain of video game research.

### 5.5. Summary of Discussion Chapter

This discussions chapter has presented a discussion of the results (Chapter 4) found for both research questions examined within this study, including a discussion of the varying interpretations and implications that may be drawn from them, placing the results of this study within the existing body of knowledge.

The discussion of results for the first research question identified several novel insights into how game events may work to elicit specific affective responses under varying conditions relevant to game play experiences more generally. The discussion of affective responses to game events demonstrates the efficacy of the method used within this study, with several valuable avenues for future research identified.

Additionally, the discussion of results for the second research question identified several novel insights into how the use of psychophysiological equipment produced several significant measurement effects that could impact the measurement and interpretation of physiological data in future studies.

There are several limitations of the results produced for both the first and second research questions, which have been discussed within this chapter. These lend themselves to future research opportunities, which are introduced and overviewed in the following chapter, alongside a summary of the research presented within this thesis.



## Chapter 6: Research Summary

### 6.1. Introduction

This chapter draws the research to a close. Previous chapters have identified the context of the research project, the specific methodological processes undertaken, and have provided insights into the two research questions posed in this thesis.

The following chapter provides a summary of the research process and the findings that have been developed. Specifically, the implications of the research findings will be presented, along with a discussion of where this research provides scope for further investigation into affective responses to video games, as well as other multimedia systems.

### 6.2. Summary of the Research

As identified in Chapters 2 and 3 of this thesis, this research project investigated affective responses to events occurring during video game play, including an exploration of measurement effects that the use of psychophysiological measures can have upon game experiences.

The recent explosion of video games as a form of entertainment has consequentially led to the rapid expansion of research investigating video games and phenomena associated, extending into peripheral fields such as psychology, Human-Computer Interaction, Information Systems, education, and even healthcare (Greitemeyer and Osswald, 2010; Lee *et al.*, 2011; Anshel, 2013; Zammitto *et al.*, 2014). However, many aspects of video games remain underexamined or poorly understood, not the least of which is the relationship between video games and emotions, especially those that occur because of interactions during game play.

The study of emotions and behaviours within video games has been traditionally dominated by subjective and qualitative measures, typically relying upon self-report surveys, interviews, and observations (Mandryk *et al.*, 2006). While these are useful instruments that can provide insights into the perceptions, experiences, and extrinsic

behaviours of people, they are generally unable to provide data on the factors that are underlying their measures due to their inability to capture data in real-time during experiences (Klasen *et al.*, 2008; Mandryk and Atkins, 2007).

Psychophysiology avoids many of the pitfalls of subjective measures due to its very nature, as it collects objective-quantitative data through passive and covert electrodes, which collect electrical signals emitted by the human body (Ijsselstein *et al.* 2008). This form of measurement can, most importantly, provide insights into experiences and affective responses in real-time, and at any point during the gameplay experience. It also provides the possibility of avoiding response bias as well as attempts at deception, and does not require participants to interpret survey questions in a consistent manner (Drachen *et al.* 2010).

However, despite the recent successes of psychophysiology and the promise it holds for the investigation of event-related phenomena occurring during video game play owing to its temporal precision and continuous measurements, very few studies attempt to examine game events in such ways (Kivikangas *et al.*, 2011; Cowley *et al.*, 2016). While there are a few notable exceptions such as studies by Ravaja *et al.* (2006), Mandryk *et al.* (2007), Weber *et al.* (2009), Vachiratamporn *et al.*, (2013), and Nogueira *et al.*, (2014), the sparsity of attempts to measure and report upon event-related affective responses occurring during video game play represents a knowledge gap within the existing body of literature.

Additionally, while the use of psychophysiological measures are relatively covert and non-invasive, the placement of electrodes upon the body and face of participants was considered to be at relatively high risk of measurement effects that may fundamentally alter the experience of video game play. Despite this increased likelihood of measurement effects, no studies were identified that attempted to examine the influence that measurement effects stemming from the use of psychophysiological equipment might have upon video game experiences or affective responses therein.

The core of this research project has therefore involved the design of a novel psychophysiological method capable of capturing data on affective responses as they occur, including sufficient contextual data to provide insights into factors that may nuance the elicitation and interpretation of observed responses.

The research problem that was investigated within this research project was:

**RP:** *How can affective responses to real-time video game events be objectively measured using a novel context-rich psychophysiological method?*

Two research questions were used to address this problem:

**RQ1:** *How do events occurring during video game play influence event-related affective responses?*

And

**RQ2:** *How does use of psychophysiological instruments influence the player experience during video game play?*

Throughout the undertaking of the methods described in Chapter 3 of this thesis, four physiological signals were used to capture objective data related to affective responses. These signals were transformed into measures of Arousal and Valence through one set of fuzzy logic rules, and then again into measures of emotions related to Fun, Excitement, Boredom, Frustration, and Challenge through a second set of fuzzy logic rules.

The cataloguing of game play data revealed ten game events that commonly occurred during game play, each of which were investigated against the measures of Arousal, Valence, and emotions described in Chapter 3. These results were examined with reference to available literature in order to determine their implications and significance to video game research. The results and literature examination of this analysis has been presented in Chapters 4 and 5 respectively.

As a result of the research process, a tentative series of implications for future research using psychophysiology to investigate emotional experiences has been developed.

### **6.2.1. Implications of the Research Findings**

The value of research is in the contributions that it can make to both its theoretical domain environment, whilst also contributing value for practical applications of the knowledge developed. The following sections of this chapter provide a discussion of both the implications that the insights developed through this research project have for the theoretical domain of video games, as well as practical implications that could be applied to either the future development or research of video games moving forward.

This study has provided valuable insights into the use of psychophysiological methods within video game contexts, including insights into the affective responses elicited through varying events occurring during video game play, and the influence of measurement effects upon the experience of participants. The contributions of the research presented here can be summarised as follows:

#### ***6.2.1.1. Identification of Real-Time Game Events and their Influences***

The psychophysiological method presented within this research project provides a way of effectively identifying game events, mechanics, or interactions that influence gameplay experiences or affective responses. Within the scope of this study, 10 game events of interest were identified and analysed, with interesting emotional responses analysed and implications for video game research and development explored. This includes the ability to measure changes in player responses to the same events or mechanics occurring at different points in time, identifying how the timing of events can be critical for the production of either a positive or a negative emotional response.

However, this method could be further applied or extended by other researchers or game developers, and can provide in-depth insights into the underlying relationships between any number of game events that occur during game play and the overall gameplay experience, including the influence that these can exert upon specific affective responses.

It is possible, for example, that game developers could use this method during gameplay testing to refine gameplay mechanics to help them develop video games that produce specific game experiences or emotional responses. It is also possible that HCI researchers could use this method to better understand how individuals respond to changes within video games, such as how changes in game pace, artwork, music, and so forth can alter gameplay experiences.

### *6.2.1.2. Identification of Real-Time Social Influences*

The method presented also provides a way of analysing the differences in affective responses and gameplay experiences between players as they play together within social contexts. Within the scope of this study, several game events were found to significantly differ depending on the perspective of the player and their compatriot.

While previous studies have identified how social interactions can produce differences in psychophysiological measures of affective responses and gameplay experiences (e.g. Klarkowski et al., 2016), none were identified that have provided a method for capturing and analysing real-time responses of both participants to the same game event.

This has several potential implications for other researchers or game developers, who could use this method to further develop understandings of how social interactions at different points in time within video games can result in significant and potentially large variations in gameplay experiences.

### *6.2.1.3. The Influence of Psychophysiology on Gameplay Experiences*

The method presented provides an examination of the influence that the presence of psychophysiological electrodes had upon the subjective gameplay experiences of participants, as well as the influence upon facial expressions and social interactions. While standard psychophysiological measures use surveys to determine whether participants report awareness of the instruments (e.g. Nacke, 2016), no previous studies to date examined the influence that such equipment had upon the actual gameplay.

This research therefore has several implications for future research, and provides preliminary evidence that psychophysiology could have significant influences upon gameplay experiences, including that participants could become distracted or introspective when electrodes are placed upon their bodies.

While the results of this study are preliminary, the evidence provided thus far warrants further exploration to confirm the effects and the develop improved understandings of how the act of using physiological measures alters the experience under examination.

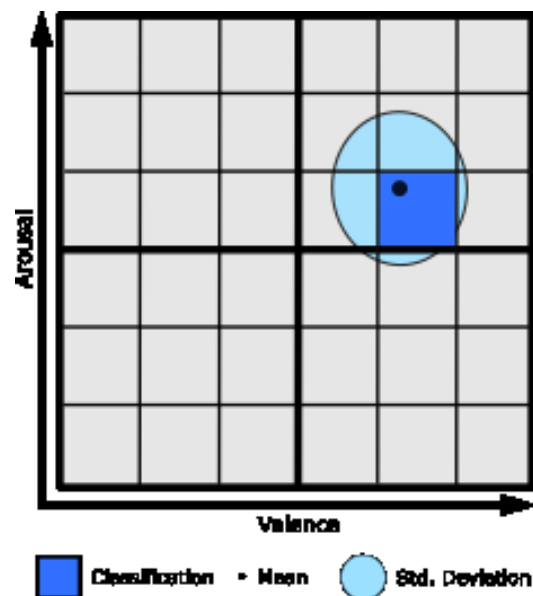
### *6.2.1.4. New Visualisation Tools for Game Events*

Some of the charts used within this project, including the Aggregated Affect Grid (e.g. Figure 97 below), the Temporal Affect Grid (e.g. Figure 91 below), and the Game Experience Affect Grid (e.g. Figure 70 below) constitute novel ways of representing game event data, which are useful for both games researchers and game developers.

#### **Aggregated Affect Grid (Mean of a Single Event):**

The Aggregated Affect Grid is a new way of providing the mean result of user affect to any single stimulus or event occurring at any point during an interactive experience. The addition of the Standard Deviation graphic for both the Arousal and Valence axes provides additional data that can assist in understanding the variances

of these experiences between users, and provides important context to the visualisation.



*Figure 97 (repeated): Aggregate Affect Grid illustrating the mean player response to a single game event, with Standard Deviation illustrated for both the Arousal and Valence axes.*

No previous research was found that represents aggregate user responses alongside standard deviation, and it is believed that this could be a valuable tool that researchers and developers alike could utilise to advance their own studies in the future.

### **Temporal Affect Grid (Changes Over Time for a Single Event):**

The Temporal Affect Grid is a new way of providing the changes that users experience over time to repetitions of the same stimulus or event during an interactive experience. This grid has been used within this thesis to illustrate aggregate data representing the average for all users (as illustrated in Figure 91 below), but can also be used to illustrate the changes of a single user over time, or of different events at the same point in time.

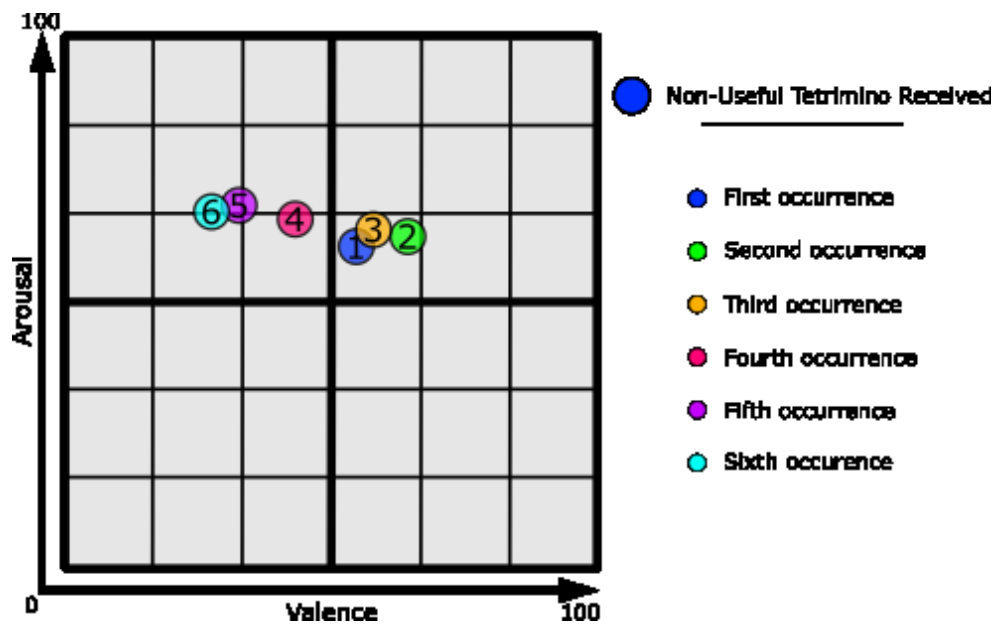


Figure 91 (repeated): Temporal Affect Grid illustrating the average affective response for users experiencing the same game event multiple times during an interactive experience.

No previous research was found that represents or compares multiple events on a single Affect Grid, and it is believed that this could be a valuable tool for improving understandings of how players experience game events or stimuli, and how events can relate to one another.

#### Game Experience Affect Grid (Mean of All Events):

The Game Experience Affect Grid is a new way of providing insights into the average experience of an interactive experience, including comparing the average overall gameplay experience against the average affective response to each game event or stimulus. This grid provides an interesting way of assessing how different events within an interactive experience influence the overall experience, and may assist researchers and developers alike to improve their understandings of how game events work together to produce an overall experience.



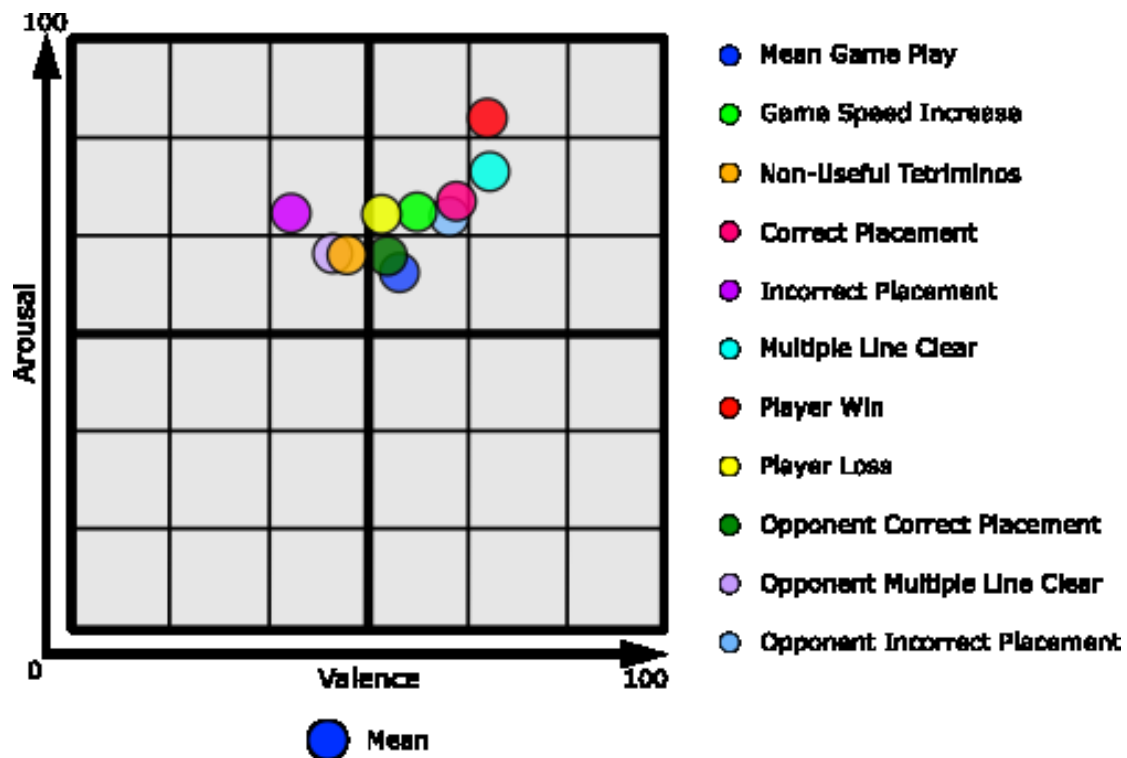


Figure 70 (repeated): Game Experience Affect Grid illustrating the average affective gameplay experience and facilitating comparisons across all game events measured within this study.

No previous research was found that represents or compares affective game experiences against its constituent events and their respective affective influences, and so it is believed that this could be a highly valuable tool that could facilitate researchers and developers alike to improve their understandings of how video game experiences are constructed through individual event experiences. Additionally, it is hoped that this could greatly improve their ability to analyse and communicate these influences.

### 6.3. Possibilities for Future Work

Given the complexity of both emotions and video games, and the value of insights that can be derived through the examination of the relationships between them, there are numerous areas in which future work should be undertaken. However, with specific reference to this research project, several valuable avenues for future work have been identified throughout the project:

- **The examination of other video games and other event types:** As this study examined only a single video game, and completed an analysis for ten events that are likely not generalizable to other games or contexts, the examination of additional video games and game events may prove a valuable extension of this project. This could include video games with more interaction between players, to examine the influence that increased interactions could have upon affective responses. This extension of the research would prove particularly valuable, as it would provide opportunities for other researchers and game developers to further test and extend the psychophysiological method presented to develop further insights into any number of within-game interactions between players and video game mechanics and events.
- **The influence of antecedents upon the elicitation and nature of affective responses to events:** As game events influence overall game experiences, it logically follows that previous game events influence the affective responses to events as they occur, with responses to events being a product of both the current context and other antecedent factors. The data collected and analysed within this project can address this point, however this would have gone beyond the scope of the research question and was therefore not included within this thesis. However, it is believed that the analysis of this data could provide more highly nuanced insights into the specific factors contributing towards event-related affective responses, and so would prove to be a valuable extension of this project. This future research would provide particularly useful for video games researchers within the HCI and psychology domains, as well as researchers within media studies and psychology more broadly, as it provides a method that has the potential to provide detailed insights and analyses into complex real-time phenomena without the requirement of isolating events from their typical context.

- **Undertake study replication and further analysis with additional instruments to improve understandings of exactly how the presence of physiological instruments changes the gameplay experience:** While the results provided within this study are promising and provide interesting insights, these should be replicated, furthered, and confirmed through additional research instruments, and within additional video games and contexts. Specifically, the Game Experience Questionnaire instrument should be substituted for more reliable subjective instruments, open-ended interviews should be undertaken, and additional objective measures such as automatic facial recordings and non-contact physiological measures should be used. If the results can be confirmed and expanded, these could provide a series of important insights that would have significant impacts upon the research community with respect to the use of physiological measures within gameplay experiences, and the situations within which psychophysiology can provide meaningful and valid results.
- **The influence of social presence and interactions upon the elicitation and nature of affective responses to events:** As this research was exploratory in nature and with limited scope and resources, it was limited in the number of experimental manipulations it could introduce and examine. While studies have examined the influence of social manipulation upon mean game play experiences, none identified have done so at the game-event level. While this study was therefore not able to do so due to its exploratory nature and limited scope, the investigation of the influence that the presence or absence of social actors has upon the experience of affective responses could prove to be a valuable extension of this project.
- **The influence of current player performance upon the elicitation and nature of affective responses to events:** Again, due to limitations in scope, this project was not able to provide an analysis of the influence that relative player performance at the time of event occurrence had upon affective

responses. The data collected and analysed within this project can address this point, however this would have gone beyond the scope of the research question and was therefore not included within this thesis. It is expected that an analysis of this data could provide additional valuable insights into how affective responses differ based upon player performance at the time of their occurrence, and so would prove to be a valuable extension of this project.

### **6.4. Concluding Reflections**

To conclude, this thesis has presented a study utilising a novel adaptation of a psychophysiological method aiming to explore affective responses to video game events as they occur during video game play. This research has provided insights into the use of psychophysiology within video game contexts, most significantly that the use of psychophysiological equipment and measures resulted in measurement effects that appear to have significantly influenced the game experience of participants.

A range of insights into specific game events is also presented alongside an analysis of the affective responses that they were found to elicit, identifying numerous factors relevant to researchers and developers. Additionally, a number of avenues for future research to extend the work of this project have been presented, which may provide additional valuable insights.

## References

- ABS. (2014). Population by Age and Sex, Regions of Australia, 2013. Available: <http://www.abs.gov.au/Ausstats/abs@.nsf/mf/3235.0>. Last accessed 4th Nov 2014.
- Ambinder, M. (2011). Biofeedback in Gameplay: How Valve Measures Physiology to Enhance Gaming Experience. Poster session presented at Information Online. Retrieved from <http://www.edge-online.com/features/gdc-2011-preview-mike-ambinder/>
- Anderson, C. a, Shibuya, A., Ihori, N., Swing, E. L., Bushman, B. J., Sakamoto, A., ... Saleem, M. (2010). Violent Video Game Effects on Aggression, Empathy, and Prosocial Behaviour in Eastern and Wesern Countries: A Meta-analytic Review. *2Psychological Bulletin*, 136(2), 151–173.
- Anshel, M. H. (2013) 'A Cognitive-Behavioral Approach for Promoting Exercise Behavior: The Disconnected Values Model.', *Journal of Sport Behavior*. Available at: <http://search.ebscohost.com/login.aspx?direct=true&andamp>.
- Ashbarry, L., Geelan, B., de Salas, K. and Lewis, I., 2016, October. Blood and Violence: Exploring the Impact of Gore in Violent Video Games. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play* (pp. 44-52). ACM.
- Astolfi, L., Cincotti, F., Mattia, D., De Vico Fallani, F., Salinari, S., Vecchiato, G., Toppi, J., Wilke, C., Doud, a, Yuan, H., He, B. and Babiloni, F. (2010) 'Simultaneous estimation of cortical activity during social interactions by using EEG hyperscannings.', *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, 2010, pp. 2814–2817. doi: 10.1109/IEMBS.2010.5626555.

Bailey, K., West, R., and Anderson, C. a. (2010). A negative association between video game experience and proactive cognitive control. *Psychophysiology*, 47(1), 34–42. doi:10.1111/j.1469-8986.2009.00925.x

Baillet, S. (2011) 'Electromagnetic Brain Mapping using MEG and EEG', *The Oxford Handbook of Social Neuroscience*, pp. 97–133.

Ballard, M., Visser, K., and Jocoy, K. (2012). Social Context and Video Game Play: Impact on Cardiovascular and Affective Responses. *Mass Communication and Society*, 15(1), 875–898. doi:10.1080/15205436.2011.632106

Banks, P.J., Tata, M.S., Bennett, P.J., Sekuler, A.B. and Gruber, A.J., 2017. Implicit Valuation of the Near-Miss is Dependent on Outcome Context. *Journal of Gambling Studies*, pp.1-17.

Bartholow, B. D., Bushman, B. J. and Sestir, M. A. (2006) 'Chronic violent video game exposure and desensitization to violence: Behavioral and event-related brain potential data', *Journal of Experimental Social Psychology*, 42(4), pp. 532–539. doi: 10.1016/j.jesp.2005.08.006.

Benedek, M. and Kaernbach, C., 2010. A continuous measure of phasic electrodermal activity. *Journal of neuroscience methods*, 190(1), pp.80-91.

Bienaimé, P. (2014). Mind-Controlled Drones Are Already A Reality. *Business Insider*. Available: <http://www.businessinsider.com.au/drones-you-can-control-with-your-mind-2014-10>. Last accessed: 27th October 2014.

BIOPAC Systems, Inc.. 2017. Products | BIOPAC. [ONLINE] Available at: <https://www.biopac.com/products/>. [Accessed 02 April 2014].

Blewitt, W. (2012) 'Exploration of Emotion Modelling Through Fuzzy Logic'. Available at: <http://hdl.handle.net/2086/6443>.

Boxtel, A. Van (2010) 'Facial EMG as a tool for inferring affective states', Proceedings of measuring behavior, 2010(Figure 3), pp. 104–108. Available at: [http://www.researchgate.net/publication/235780510\\_Facial\\_EMG\\_as\\_a\\_tool\\_for\\_inferring\\_affective\\_states/file/32bfe513736978ba42.pdf](http://www.researchgate.net/publication/235780510_Facial_EMG_as_a_tool_for_inferring_affective_states/file/32bfe513736978ba42.pdf).

Braddon-Mitchell, D. and Jackson, F., 2006. Philosophy of mind and cognition: An introduction.

Bradley, M. and Lang, P. J. (1994) 'Measuring Emotion: The Self-Assessment Semantic Differential Manikin and the', Journal of Behavior Therapy and Experimental Psychiatry, 25(1), pp. 49–59. doi: 10.1016/0005-7916(94)90063-9.

Brand, J. E., Lorentz, P. and Mathew, T. (2014) Digital Australia 2014. Available at <http://www.igea.net/wp-content/uploads/2015/07/Digital-Australia-2016-DA16-Final.pdf>

Brockmyer, J. H., Fox, C. M., Curtiss, K. A., McBroom, E., Burkhart, K. M., and Pidruzny, J. N. (2009). The development of the Game Engagement Questionnaire: A measure of engagement in video game-playing. Journal of Experimental Social Psychology. Retrieved from <http://www.sciencedirect.com.ezproxy.utas.edu.au/science/article/pii/S0022103109000444?np=y>

Brühlmann, F. and Schmid, G.M., 2015, April. How to measure the game experience?: Analysis of the factor structure of two questionnaires. In Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (pp. 1181-1186). ACM.

Brunner, C., Delorme, A. and Makeig, S. (2013) 'EEGLAB an open source MATLAB toolbox for electrophysiological research', *Biomed Tech*, 58(1), pp. 24–25. doi: 10.1515/bmt-2013-4.

Burgess, A. P. (2013) 'On the interpretation of synchronization in EEG hyperscanning studies: a cautionary note.', *Frontiers in human neuroscience*, 7(December), p. 881. doi: 10.3389/fnhum.2013.00881.

Burrell, G. and Morgan, G., 1979. *Sociological paradigms and organisational analysis* (Vol. 248). London: Heinemann.

Cacioppo, J. T. and Tassinary, L. G. (1990) 'Inferring psychological significance from physiological signals.', *American Psychologist*, 45(1), pp. 16–28. doi: 10.1037/0003-066X.45.1.16.

Cacioppo, J., Tassinary, L. G. and Berntson, G. G. (2007) *The Handbook of Psychophysiology*. 3rd edn. New York, New York, USA: Cambridge University Press.

Calvo, R. A. and D'Mello, S. (2010) 'Affect detection: An interdisciplinary review of models, methods, and their applications', *IEEE Transactions on Affective Computing*, 1(1), pp. 18–37. doi: 10.1109/T-AFFC.2010.1.

Ceaparu, I., Lazar, J., Bessiere, K., Robinson, J. and Shneiderman, B., 2004. Determining causes and severity of end-user frustration. *International journal of human-computer interaction*, 17(3), pp.333-356.

Chen, J. (2007) 'Flow in games (and everything else)', *Communications of the ACM*. ACM Request Permissions, 50(4), p. 31. Available at: <http://portal.acm.org/citation.cfm?doid=1232743.1232769>.



Christoph, K., Dorothée, H. and Peter, V. (2009) 'The video game experience as "true" identification: A theory of enjoyable alterations of players' self-perception', *Communication Theory*, 19(4), pp. 351–373. doi: 10.1111/j.1468-2885.2009.01347.x.

Clark, M. (2012) *Measuring the Cooperative Play Experience from Electroencephalographic Readings*.

Cole, S. W., Yoo, D. J., and Knutson, B. (2012). Interactivity and reward-related neural activation during a serious videogame. *PLoS ONE*, 7(3), 1–10. Retrieved from <http://dx.plos.org/10.1371/journal.pone.0033909.g005>

Cowley, B., Fantato, M., Jennett, C., Ruskov, M. and Ravaja, N. (2014) 'Learning When Serious : Psychophysiological Evaluation of a Technology- Enhanced Learning Game', *Educational Technology and Society*, 17(1), pp. 3–16.

Cowley, B., Filetti, M., Lukander, K., Torniainen, J., Henelius, A., Ahonen, L., Barral, O., Kosunen, I., Valtonen, T., Huotilainen, M., Ravaja, N. and Jacucci, G. (2016) 'The Psychophysiology Primer : A Guide to Methods and a Broad Review with a Focus on Human – Computer Interaction The Psychophysiology Primer : A Guide to Methods and a Broad Review with a Focus on Human-Computer Interaction', *Foundations and Trends in Human-Computer Interaction*, 9(3–4), pp. 151–308. doi: 10.1561/11000000065.

Crossan, F., 2003. Research philosophy: towards an understanding. *Nurse researcher*, 11(1), pp.46-55.

Crutzen, R., van 't Riet, J. and Short, C. E. (2015) 'Enjoyment: A Conceptual Exploration and Overview of Experimental Evidence in the Context of Games for Health.', *Games for health journal*, 5(1), pp. 15–20. doi: 10.1089/g4h.2015.0059.

Csikszentmihalyi, M., 1975. Play and intrinsic rewards. *Journal of humanistic psychology*.

de Gialdino Irene, V., 2011. Ontological and epistemological foundations of qualitative research [85 paragraphs]. In *Forum Qualitative Sozialforschung/Forum Qual Soc Res* (Vol. 10, No. 2).

Decety, J. and Cacioppo, J.T., 2010. Problems and Prospects in Social Neuroscience. *Japanese Journal of Physiological Psychology and Psychophysiology*, 28(1), pp.5-16.

Decety, J., Jackson, P. L., Sommerville, J. a, Chaminade, T., and Meltzoff, A. N. (2004). The neural bases of cooperation and competition: an fMRI investigation. *NeuroImage*, 23(2), 744–51. doi:10.1016/j.neuroimage.2004.05.025

Deci, E.L. and Ryan, R.M., 2008. Self-determination theory: A macrotheory of human motivation, development, and health. *Canadian psychology/Psychologie canadienne*, 49(3), p.182.

Delorme, A. and Makeig, S. (2004) 'EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis', *Journal of Neuroscience Methods*, 134(1), pp. 9–21. doi: 10.1016/j.jneumeth.2003.10.009.

Delsys, Inc.. 2017. EMGworks Software | Delsys, Inc.. [ONLINE] Available at: <http://www.delsys.com/product/emgworks-software/>. [Accessed 14 May 2015].

Denzin, N.K. and Lincoln, Y.S., 1998. [Vol. 1]: The landscape of qualitative research: theories and issues. Thousand Oaks [etc.]: Sage.

Deterding, S., 2012. Gamification: designing for motivation. *interactions*, 19(4), pp.14-17.

Deterding, S., Dixon, D., Khaled, R. and Nacke, L., 2011, September. From game design elements to gamefulness: defining gamification. In Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments (pp. 9-15). ACM.

DigiFit. (2014). DigiFit. Available from: <http://www.digifit.com/>. Last accessed 17th August 2014.

Dixon, M.J., Harrigan, K.A., Jarick, M., MacLaren, V., Fugelsang, J.A. and Sheepy, E., 2011. Psychophysiological arousal signatures of near-misses in slot machine play. *International Gambling Studies*, 11(3), pp.393-407.

Dolcos, F., Iordan, A. D., and Dolcos, S. (2011). Neural correlates of emotion–cognition interactions: A review of evidence from brain imaging investigations. *Journal of Cognitive Psychology*, 23(6), 669–694. Retrieved from <http://www.tandfonline.com/doi/abs/10.1080/20445911.2011.594433>

Drachen, A., Nacke, L., Yannakakis, G. and Pedersen, A. L. (2010) ‘Correlation between heart rate, electrodermal activity and player experience in first-person shooter games’, ... on Video Games. Available at: <http://dl.acm.org/citation.cfm?id=1836143> (Accessed: 17 September 2014).

Drachen, A., Nacke, L.E., Yannakakis, G. and Pedersen, A.L., 2010, July. Correlation between heart rate, electrodermal activity and player experience in first-person shooter games. In Proceedings of the 5th ACM SIGGRAPH Symposium on Video Games (pp. 49-54). ACM.

Duvinage, M., Castermans, T., Petieau, M., Hoellinger, T., Cheron, G., and Dutoit, T. (2013). Performance of the Emotiv Epoc headset for P300-based applications. *Biomedical Engineering Online*, 12, 56. doi:10.1186/1475-925X-12-56

Dworak, M., Schierl, T., Bruns, T., and Strüder, H. K. (2007). Impact of singular excessive computer game and television exposure on sleep patterns and memory performance of school-aged children. *Pediatrics*, 120(5), 978–85.

doi:10.1542/peds.2007-0476

Dye, M. W. G., Green, C. S., and Bavelier, D. (2009). Increasing Speed of Processing With Action Video Games. *Current Directions in Psychological Science*, 18(6), 321–326. doi:10.1111/j.1467-8721.2009.01660.

EEGLAB - Matlab software for signal processing of electrophysiological brain data . 2017. EEGLAB - Matlab software for signal processing of electrophysiological brain data . [ONLINE] Available

at: <http://cognitrn.psych.indiana.edu/busey/temp/eeglabtutorial4.301/>. [Accessed 19 July 2015].

Ekman, P. (2014). Darwin’s claim of universals in facial expression not challenged.

Paul Ekman Group. Available from:

<http://www.paulekman.com/uncategorized/darwins-claim-universals-facial-expression-challenged/>. Last accessed: 18th July 2014.

Ellis, G.D., Voelkl, J.E. and Morris, C., 1994. Measurement and analysis issues with explanation of variance in daily experience using the flow model. *Journal of leisure research*, 26(4), p.337.

El-Nasr, M.S., Aghabeigi, B., Milam, D., Erfani, M., Lameman, B., Maygoli, H. and Mah, S. (2010). Understanding and evaluating cooperative games. In *Proceedings of the 28th International Conference on Human Factors in Computing Systems* (pp. 253-262). Atlanta, GA.

EMGLAB. 2017. EMGLAB. [ONLINE] Available at: <http://www.emglab.net/>. [Accessed 09 July 2015].

Emotiv. 2017. EMOTIV Epoc - 14 Channel Wireless EEG Headset. [ONLINE] Available at: <https://www.emotiv.com/epoc/>. [Accessed 19 Jan 2015].

Ermi, L. and Mäyrä, F. (2005) 'Fundamental Components of the Gameplay Experience : Analysing Immersion', in Digital Games Research Association 2nd International Conference, pp. 15–27.

Fatourechi, M., Ward, R. K., Bashashati, A. and Birch, G. E. (2007) 'A survey of signal processing algorithms in brain-computer interfaces based on electrical brain signals', *Journal of Neural Engineering*, 4(2), pp. R32-57. doi: 10.1088/1741-2560/4/2/R03.

Felmingham, K.L., Bryant, R.A. and Gordon, E., 2003. Processing angry and neutral faces in post-traumatic stress disorder: an event-related potentials study. *Neuroreport*, 14(5), pp.777-780.

Ferguson, J. (2007). The Good, The Bad and the Ugly: A Meta-analytic Review of Positive and Negative Effects of Violent Video Games. *Psychiatry Quarterly*, 78, 309–316.

Ferro, L., Walz, S., and Greuter, S. (2013). Towards personalised, gamified systems: an investigation into game design, personality and player typologies. ... of The 9th Australasian Conference on ..., 1–6. Retrieved from <http://dl.acm.org/citation.cfm?id=2513024>

Foster, C., Green, M. A., Snyder, A. C., and Thompson, N. N. (1993). Physiological responses during simulated competition. *Medical Science Sports Exercise*, 25(7), 877-882.

Fowles, D.C., 1980. The three arousal model: Implications of Gray's two-factor learning theory for heart rate, electrodermal activity, and psychopathy. *Psychophysiology*, 17(2), pp.87-104.

Frattaroli, J., 2006. Experimental disclosure and its moderators: a meta-analysis. *Psychological bulletin*, 132(6), p.823.

Gago, D., and Almeida, R. (2013). Effects of pleasant visual stimulation on attention, working memory, and anxiety in college students. *Psychology and Neuroscience*, 6(3), 351–355. Retrieved from [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S1983-32882013000300012&lng=pt&dnrm=iso&dtlng=en](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1983-32882013000300012&lng=pt&dnrm=iso&dtlng=en)

Gajadhar, B., de Kort, Y., and IJsselstein, W. (2008). Shared fun is doubled fun: Player enjoyment as a function of social setting. *Fun and Games*, 5294, 106-117. doi: 10.1007/978-3-540-88322-7

Ghosh, D. and Vogt, A., 2012. Outliers: An evaluation of methodologies. In *Joint Statistical Meetings* (pp. 3455-3460). San Diego, CA: American Statistical Association.

Girden, E.R., 1992. *ANOVA: Repeated measures* (No. 84). Sage.

GLNeuroTechnologies. (2014). The BioRadio Wireless Physiological Monitor. Available: <http://glneurotech.com/bioradio/bioradio-wireless-physiological-monitor/>. Last accessed 1st December 2014.

Gorham, G. (1994) 'Mind-body dualism and the Harvey-Descartes controversy.', *Journal of the history of ideas*, 55(2), pp. 211–234. doi: 10.2307/2709897.

Greenberg, B.D., Gabriels, L.A., Malone, D.A., Rezai, A.R., Friehs, G.M., Okun, M.S., Shapira, N.A., Foote, K.D., Cosyns, P.R., Kubu, C.S. and Malloy, P.F., 2010. Deep brain

stimulation of the ventral internal capsule/ventral striatum for obsessive-compulsive disorder: worldwide experience. *Molecular psychiatry*, 15(1), pp.64-79.

Greenhouse, S.W. and Geisser, S., 1959. On methods in the analysis of profile data. *Psychometrika*, 24(2), pp.95-112.

Greitemeyer, T. and Osswald, S. (2010) 'Effects of Prosocial Video Games on Prosocial Behaviour', *Journal of Personality and Social Psychology*, 98(2), pp. 211–221.

Gross, J.J., 2002. Emotion regulation: Affective, cognitive, and social consequences. *Psychophysiology*, 39(3), pp.281-291.

Guba, E.G. ed., 1990. *The paradigm dialog*. Sage Publications.

Hall Jr, D.P. and Jansen, J.A., 1995. Stress and arousal in deployment of a combat support hospital. *Military medicine*, 160(11), pp.581-583.

Hall, J. (2011). *Guyton and Hall textbook of medical physiology* (12th ed. ed.). Philadelphia, Pa.: Saunders/Elsevier. p. 157.

Hamari, J., Koivisto, J. and Sarsa, H., 2014, January. Does gamification work?--a literature review of empirical studies on gamification. In *System Sciences (HICSS), 2014 47th Hawaii International Conference on* (pp. 3025-3034). IEEE.

Han, D. H., Kim, Y. S., Lee, Y. S., and Min, K. J. (2010). Changes in cue-induced, prefrontal cortex activity with video-game play. *Review of Educational Research*, 13(6), 655–661. Retrieved from <http://www.liebertonline.com/doi/abs/10.1089/cyber.2009.0327>

Havranek, M., Langer, N., Cheetham, M., and Jäncke, L. (2012). Perspective and agency during video gaming influences spatial presence experience and brain activation patterns. *Behavioral and Brain Functions*, 8(1), 34-38. doi:10.1186/1744-9081-8-34

Hernandez, J., Riobo, I., Rozga, A., Abowd, G. D., and Picard, R. W. (2014). Using Electrodermal Activity to Recognize Ease of Engagement in Children during Social Interactions. In *Ubiquitous Computing*.

Hilgard, J., Engelhardt, C. R., and Bartholow, B. D. (2013). Individual differences in motives, preferences, and pathology in video games: the gaming attitudes, motives, and experiences scales (GAMES). *Frontiers in Psychology*, 4, 608. doi:10.3389/fpsyg.2013.00608

Homma, I., and Masaoka, Y. (2008). Breathing rhythms and emotions. *Exploratory Physiology*, 93(1), 1011–1021.

Howell, D.C., 2012. *Statistical methods for psychology*. Cengage Learning.

IBM SPSS Statistics. 2017. IBM SPSS Statistics. [ONLINE] Available at: <https://www.ibm.com/us-en/marketplace/spss-statistics>. [Accessed 04 Feb 2016].

Ijsselsteijn, W., Hoogen, W. Van Den, Klimmt, C., Kort, Y. De, Lindley, C., Mathiak, K., ... Vorderer, P. (2008). Measuring the Experience of Digital Game Enjoyment. In A. J. Spink, M. R. Ballintjin, N. D. Bogers, F. Grieco, L. W. S. Loijens, L. P. J. J. Noldus, ... P. H. Zimmerman (Eds.), *Proceedings of Measuring Behavior* (Vol. 2008, pp. 88–90). Maastricht, The Netherlands.

Ivory, J. D. and Kalyanaraman, S. (2007) 'The effects of technological advancement and violent content in video games on players' feelings of presence, involvement,



physiological arousal, and aggression', *Journal of Communication*, 57(3), pp. 532–555. doi: 10.1111/j.1460-2466.2007.00356.x.

Jakymin, A. K. and Harris, L. M. (2012) 'Self-focused attention and social anxiety', *Australian Journal of Psychology*, 64(2), pp. 61–67. doi: 10.1111/j.1742-9536.2011.00027.x.

Janis, J.E., Ghavami, A., Lemmon, J.A., Leedy, J.E. and Guyuron, B., 2007. Anatomy of the corrugator supercilii muscle: Part I. Corrugator topography. *Plastic and reconstructive surgery*, 120(6), pp.1647-1653.

Jarvela, S., Kivikangas, J. M., Katsyri, J., Henttonen, P. and Ravaja, N. (2014) 'Physiological Linkage of Dyadic Gaming Experience', *Simulation and Gaming*, 45(1), pp. 24–40.

Jegers, K., 2007. Pervasive game flow: understanding player enjoyment in pervasive gaming. *Computers in Entertainment (CIE)*, 5(1), p.9.

Jin, S.-A. A. (2012) "'Toward Integrative Models of Flow": Effects of Performance, Skill, Challenge, Playfulness, and Presence on Flow in Video Games', *Journal of Broadcasting and Electronic Media*, 56(2), pp. 169–186. Available at: <http://www.tandfonline.com/doi/abs/10.1080/08838151.2012.678516>.

Johnson, D., and Gardner, J. (2010). Personality, Motivation and Video Games. In *OZCHI 2010 Proceedings*, Brisbane, QLD. 276-279.

Johnson, R. and Onwuegbuzie, A. (2004) 'Mixed methods research: A research paradigm whose time has come', *Educational Researcher*, 33(7), pp. 14–26. doi: 10.2307/3700093.

Jones, R. and Rowe, M. (2015) 'A Guide for Analysing Electrodermal Activity ( EDA ) and Skin Conductance Responses ( SCRs ) for Psychological Experiments'.

Kaplan, B. and Duchon, D., 1988. Combining qualitative and quantitative methods in information systems research: a case study. *MIS quarterly*, pp.571-586.

Kassam, K. S. and Mendes, W. B. (2013) 'The Effects of Measuring Emotion: Physiological Reactions to Emotional Situations Depend on whether Someone Is Asking', *PLoS ONE*, 8(6), pp. 6–13. doi: 10.1371/journal.pone.0064959.

Keller, J., Bless, H., Blomann, F. and Kleinböhl, D., 2011. Physiological aspects of flow experiences: Skills-demand-compatibility effects on heart rate variability and salivary cortisol. *Journal of Experimental Social Psychology*, 47(4), pp.849-852.

Kirman, B. (2010) 'Emergence and Playfulness in Social Games', 44(0), pp. 71–77.

Kirman, B., Paavilainen, J., Bjork, S., Rao, V. and Deterding, S. (2011) 'Social Game Studies at CHI 2011', pp. 1–4. Available at:

[http://dl.dropboxusercontent.com/u/220532/3.Social Games Workshop Abstract.pdf](http://dl.dropboxusercontent.com/u/220532/3.Social%20Games%20Workshop%20Abstract.pdf).

Kirmizi-Alsan, E., Bayraktaroglu, Z., Gurvit, H., Keskin, Y. H., Emre, M. and Demiralp, T. (2006) 'Comparative analysis of event-related potentials during Go/NoGo and CPT: decomposition of electrophysiological markers of response inhibition and sustained attention.', *Brain research*, 1104(1), pp. 114–28. doi: 10.1016/j.brainres.2006.03.010.

Kivikangas, J. M., Ekman, I., Chanel, G., Järvelä, S., Cowley, B., Henttonen, P. and Ravaja, N. (2011) 'Review on psychophysiological methods in game research', *Journal of Gaming and Virtual Worlds*, 3(3), pp. 181–199.

Klabunde, R. (2013). Myocardial Oxygen Demand. Cardiovascular Physiology Concepts. Available: <http://cvphysiology.com/CAD/CAD003.htm>. Last accessed: 28th June 2014.

Klarkowski, M., Johnson, D., Wyeth, P., Phillips, C. and Smith, S. (2016) 'Psychophysiology of Challenge in Play : EDA and Self-Reported Arousal', (June). doi: 10.1145/2851581.2892485.

Klasen, M., Weber, R., Kircher, T. T. J., Mathiak, K. A. and Mathiak, K. (2012) 'Neural contributions to flow experience during video game playing', *Social Cognitive and Affective Neuroscience*, 7(4), pp. 485–495. Available at: <http://scan.oxfordjournals.org/lookup/doi/10.1093/scan/nsr021>.

Klasen, M., Weber, R., Mathiak, K. A., Zvyagintsev, M., and Mathiak, K. (2008). Think Aloud-method during fMRI to determine neuronal correlates of subjective experience of video game playing. In *Proceedings of Measuring Behavior* (pp. 95–67). Maastricht, The Netherlands.

Klimmt, C., Blake, C., Hefner, D., Vorderer, P. and Roth, C. (2009) 'Player Performance , Satisfaction , and Video Game', *Ifip International Federation For Information Processing*, 5709 LNCS, pp. 1–12. Available at: [http://dx.doi.org/10.1007/978-3-642-04052-8\\_1](http://dx.doi.org/10.1007/978-3-642-04052-8_1).

Kneer, J., Elson, M. and Knapp, F. (2016) 'Fight fire with rainbows: The effects of displayed violence, difficulty, and performance in digital games on affect, aggression, and physiological arousal', *Computers in Human Behavior*, 54(JANUARY), pp. 142–148. doi: 10.1016/j.chb.2015.07.034.

Koeffel, C., Hochleitner, W., Leitner, J., Haller, M., Geven, A. and Tscheligi, M. (2010) 'Using Heuristics to Evaluate the Overall User Experience of Video Games and

Advanced Interaction Games', *Evaluating User Experience in Games: Concepts and Methods*, (January), pp. 233–256. doi: 10.1007/978-1-84882-963-3.

Koelstra, S., and Muhl, C. (2012). Deap: A database for emotion analysis; using physiological signals. *Affective Computing*, ..., 1–15. Retrieved from [http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=5871728](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5871728)

Koelstra, S., Muhl, C., Soleymani, M., Lee, J.S., Yazdani, A., Ebrahimi, T., Pun, T., Nijholt, A. and Patras, I., 2012. Deap: A database for emotion analysis; using physiological signals. *IEEE Transactions on Affective Computing*, 3(1), pp.18-31.

Kokil, U., 2013. Investigating Players' Affective States in an Interactive Environment. *ThinkMind//ACHI*.

Kong, D. and Theodore, M., 2011. Competitive video games and social capital: New frontiers of community formation. *America*.

Kring, A. and Sloan, D. (1991) 'The facial expression coding system (FACES): A users guide', Unpublished manuscript. doi: citeulike-article-id:3171485.

Kuhn, S., Romanowski, A., Schilling, C., *et al.* (2011). The neural basis of video gaming. *Translational Psychiatry*, 1:e53.

Kumar, M., Arndt, A., Kreuzfeld, S., Thurow, K., Stoll, N. and Stoll, R. (2008) 'Fuzzy techniques for subjective workload-score modeling under uncertainties', *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 38(6), pp. 1449–1464. doi: 10.1109/TSMCB.2008.927712.

Laerd Statistics (2015). One-way repeated measures ANOVA using SPSS Statistics. Statistical tutorials and software guides. Retrieved from <https://statistics.laerd.com/>

Lambie, J.A. and Marcel, A.J., 2002. Consciousness and the varieties of emotion experience: a theoretical framework. *Psychological review*, 109(2), p.219.

Lang, F. R., John, D., Lüdtke, O., Schupp, J., and Wagner, G. G. (2011). Short assessment of the Big Five: robust across survey methods except telephone interviewing. *Behavior Research Methods*, 43(2), 548–67. doi:10.3758/s13428-011-0066-z

Lang, P.J. (1995). The Emotion Probe. *Studies of motivation and attention. American Psychologist*, 50(5), 372-385. doi: 10.1037/0003-066X.50.5.372

Lang, P.J., 1985. The cognitive psychophysiology of emotion: Fear and anxiety.

Lang, P.J., Greenwald, M.K., Bradley, M.M. and Hamm, A.O., 1993. Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology*, 30(3), pp.261-273.

Larsen, J.T., Berntson, G.C., Poehlmann, K.M., Ito, T.A. and Cacioppo, J.T. (2000). The Psychophysiology of Emotion. In Lewis, M. and Haviland, J.M (Eds.), *Handbook of Emotions* (pp. 173 – 191). doi: 10.1111/j.1469-8986.1989.tb01959.x

Ledalab. 2017. Ledalab. [ONLINE] Available at: <http://www.ledalab.de/>. [Accessed 17 May 2015].

Lee, K. M., Jeong, E. J., Park, N. and Ryu, S. (2011) 'Effects of Interactivity in Educational Games: A Mediating Role of Social Presence on Learning Outcomes', *International Journal of Human-Computer Interaction*, 27(7), pp. 620–633. Available at: <http://www.tandfonline.com/doi/abs/10.1080/10447318.2011.555302>.

Left 4 Dead 2 [Computer software]. (2009). Bellevue, WA: Valve Corporation. 60

Levillain, F., Orero, J. O., Rifqi, M. and Bouchon-meunier, B. (2010) 'Characterizing Player ' s Experience From Physiological Signals Using Fuzzy Decision Trees', *Computational Intelligence*, pp. 75–82.

Lieberman, M. D., Eisenberger, N. I., Crockett, M. J., Tom, S. M., Pfeifer, J. H. and Way, B. M. (2007) 'Putting feelings into words: Affect labeling disrupts amygdala activity in response to affective stimuli: Research article', *Psychological Science*, 18(5), pp. 421–428. doi: 10.1111/j.1467-9280.2007.01916.x.

Lim, S. and Reeves, B. (2010). Computer agents versus avatars: Responses to interactive game characters controlled by a computer or other player. *International Journal of Human- Computer Studies*, 68(1), 57-68. doi: 10.1016/j.ijhcs.2009.09.008

Liu, C.-C. and Chang, I.-C. (2012) 'Measuring The Flow Experience Of Players Playing Online Games', in *Proceedings of the PACIS 2012 Conference*, pp. 1–16. Available at: [papers3://publication/uuid/F1C24B9E-E001-4708-A2C6-3EE2224A0944](http://papers3://publication/uuid/F1C24B9E-E001-4708-A2C6-3EE2224A0944).

Liu, Y., Sourina, O. and Nguyen, M. K. (2011) 'Real-time EEG-based emotion recognition and its applications', *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6670 LNCS, pp. 256–277. doi: 10.1007/978-3-642-22336-5\_13.

Lyons, E. J. (2015) 'Cultivating Engagement and Enjoyment in Exergames Using Feedback, Challenge, and Rewards', *Games for Health Journal*, 4(1), pp. 12–18. doi: 10.1089/g4h.2014.0072.

Ma, W. (2014). ING Direct research reveals how Australians spend their monthly salaries. Available: <http://www.news.com.au/finance/money/ing-direct-research-reveals-how-australians-spend-their-monthly-salaries/story-e6frfmcr-1226963930282>. Last accessed 17th July 2014.

Mandryk, R. L. and Inkpen, K. M. (2004) 'Physiological indicators for the evaluation of co-located collaborative play', Proceedings of the 2004 ACM conference on Computer supported cooperative work - CSCW '04. New York, New York, USA: ACM Press, p. 102. doi: 10.1145/1031607.1031625.

Mandryk, R. L., and Atkins, M. S. (2007). A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies. *International Journal of Human-Computer Studies*, 65(4), 329–347. doi:10.1016/j.ijhcs.2006.11.011

Mandryk, R. L., Atkins, M. S. and Inkpen, K. M. (2006) 'A continuous and objective evaluation of emotional experience with interactive play environments', Proceedings of the SIGCHI conference on Human Factors in computing systems - CHI '06. New York, New York, USA: ACM Press, p. 1027. doi: 10.1145/1124772.1124926.

Mandryk, R.L. and Nacke, L.E., 2016. Biometrics in gaming and entertainment technologies (pp. 197-224).

Marker, A. M. and Staiano, A. E. (2015) 'Better Together: Outcomes of Cooperation Versus Competition in Social Exergaming.', *Games for health journal*, 4(1), pp. 25–30. doi: 10.1089/g4h.2014.0066.

Matiko, J. W., Beeby, S. P. and John, T. (2014) 'Fuzzy logic based emotion classification', 2014 IEEE International Conference on Acoustic, Speech and Signal Processing, (MAY 2014), pp. 4422–4426. doi: 10.1109/ICASSP.2014.6854431.

Maxwell, S.E. and Delaney, H.D., 2004. Designing experiments and analyzing data: A model comparison perspective (Vol. 1). Psychology Press.

Mayers, A., 2013. Introduction to Statistics and SPSS in Psychology. Pearson.

McGill, K. C., Lateva, Z. C. and Marateb, H. R. (2005) 'EMGLAB: An interactive EMG decomposition program', *Journal of Neuroscience Methods*, 149(2), pp. 121–133. doi: 10.1016/j.jneumeth.2005.05.015.

Mellecker, R., Lyons, E. J., and Baranowski, T. (2013). Disentangling Fun and Enjoyment in Exergames Using an Expanded Design, Play, Experience Framework: A Narrative Review. *Games for Health Journal*, 2(3), 142–149. Retrieved from <http://online.liebertpub.com/doi/abs/10.1089/g4h.2013.0022>

Metacritic. (2012). Tetris Party Deluxe for Wii. Available: <http://www.metacritic.com/game/wii/tetris-party-deluxe>. Last accessed: 2nd August 2014.

Miller, J.G., 1978. Living systems.

Mingers, J., 2001. Combining IS research methods: towards a pluralist methodology. *Information systems research*, 12(3), pp.240-259.

Morris, J. D. (1995) 'OBSERVATIONS: SAM: The Self-Assessment Manikin - An Efficient Cross-Cultural Measurement of Emotional Response', *Journal of Advertising Research*, 35(6), pp. 63–68.

Morris, J.D., Bradley, M.M. and Karrh, J.A., 1994. Assessing Affective Response to Television Advertising Using the Self-Assessment Manikin (SAM). Working paper, University of Florida, College of Journalism and Communications.

Mosher, J. and Baillet, S. (2005) 'Brainstorm electromagnetic imaging software', *International ...*, 7(2), pp. 189–190. Available at: [http://neuroimage.usc.edu/paperspdf/Mosher\\_NSFI\\_2005.pdf](http://neuroimage.usc.edu/paperspdf/Mosher_NSFI_2005.pdf).



Muender, T., Miller, M. K., Birk, M. V and Mandryk, R. L. (2016) 'Extracting Heart Rate from Videos of Online Participants', Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 4562–4567.

Myers, M.D., 1997. Qualitative research in information systems. Management Information Systems Quarterly, 21(2), pp.241-242.

Nacke, L., 2009. Affective ludology: Scientific measurement of user experience in interactive entertainment (Doctoral dissertation, Blekinge Institute of Technology).

Nacke, L.E., 2010, May. Wiimote vs. controller: electroencephalographic measurement of affective gameplay interaction. In Proceedings of the International Academic Conference on the Future of Game Design and Technology (pp. 159-166). ACM.

Nacke, L.E., 2013. An introduction to physiological player metrics for evaluating games. In Game Analytics (pp. 585-619). Springer London.

Nacke, L.E., 2015. Games user research and physiological game evaluation. In Game user experience evaluation (pp. 63-86). Springer International Publishing.

Nacke, L. and Lindley, C. A. (2009) 'Affective Ludology , Flow and Immersion in a First- Person Shooter : Measurement of Player Experience', The Journal of the Canadian Game Studies Association, 3(5).

Nacke, L. E., and Mandryk, R. L. (2011). Designing Affective Games with Physiological Input. Fun and Games.

Nacke, L. E., Nacke, A. and Lindley, C. A. (2009) 'Brain Training for Silver Gamers: Effects of Age and Game Form on Effectiveness, Efficiency, Self-Assessment, and

Gameplay Experience', *CyberPsychology and Behavior*, 12(5), pp. 493–499. doi: 10.1089/cpb.2009.0013.

Nacke, L. E., Stellmach, S., and Lindley, C. A. (2011). Electroencephalographic Assessment of Player Experience: A Pilot Study in Affective Ludology. *Simulation and Gaming*, 42(5), 632–655. doi:10.1177/1046878110378140

Nacke, L., Ambinder, M., Canossa, a and Mandryk, R. (2009) 'Game Metrics and Biometrics: The Future of Player Experience Research', *Future ...*, (July 2016), pp. 2–5. Available at: [https://www.netlearning2002.org/fou/forskinfo.nsf/all/301f20d8b293f29ec12575c50058ad46/\\$file/Nacke-et-al-Game-Metrics-Panel.pdf%5Cnpapers2://publication/uuid/FC317994-10B4-475B-9243-8E6C8A2CA586](https://www.netlearning2002.org/fou/forskinfo.nsf/all/301f20d8b293f29ec12575c50058ad46/$file/Nacke-et-al-Game-Metrics-Panel.pdf%5Cnpapers2://publication/uuid/FC317994-10B4-475B-9243-8E6C8A2CA586).

Nacke, L., Drachen, A., and Göbel, S. (2010). Methods for Evaluating Gameplay Experience in a Serious Gaming Context. *International Journal of Computer Science in Sport*, 9(2), 40–51.

Nagamitsu S, Nagano M, Yamashita Y, *et al.* (2006). Prefrontal cerebral blood volume patterns while playing video games—a near-infrared spectroscopy study. *Brain Development*, 28(1), 315–321.

Nogueira, P. A., Aguiar, R., Rodrigues, R., Oliveira, E. and Nacke, L. E. (2014) 'Fuzzy Affective Player Models: A Physiology-Based Hierarchical Clustering Method', *Proceedings of the Tenth Artificial Intelligence and Interactive Digital Entertainment Conference, (Aiide)*, pp. 132–138. Available at: <http://www.aaai.org/ocs/index.php/AIIDE/AIIDE14/paper/view/8947>.

Norman, K.L., 2013. Geq (game engagement/experience questionnaire): a review of two papers. *Interacting with Computers*, 25(4), pp.278-283.

Olofsson, J.K., Nordin, S., Sequeira, H. and Polich, J., 2008. Affective picture processing: an integrative review of ERP findings. *Biological psychology*, 77(3), pp.247-265.

OMRON. (2014). Expression Estimation. Available from: <http://www.omron.com/ecb/products/mobile/okao07.html>. Last accessed: 23rd August 2014.

Open Broadcaster Software | Home. 2017. Open Broadcaster Software | Home. [ONLINE] Available at: <https://obsproject.com/>. [Accessed 05 April 2015].

Orero, J. O., Levillain, F., Damez-Fontaine, M., Rifqi, M. and Bouchon-Meunier, B. (2010) 'Assessing Gameplay Emotions from Physiological signals - a Fuzzy Decision

Trees Based Model', *Kansei Engineering and Emotion Research International Conference 2010*, (May), pp. 1684–1693.

Orlikowski, W.J. and Baroudi, J.J., 1991. Studying information technology in organizations: Research approaches and assumptions. *Information systems research*, 2(1), pp.1-28.

Pattyn, N., Neyt, X., Henderickx, D. and Soetens, E. (2008) 'Psychophysiological investigation of vigilance decrement: Boredom or cognitive fatigue?', *Physiology and Behavior*, 93(1–2), pp. 369–378. doi: 10.1016/j.physbeh.2007.09.016.

Peng, W., and Hsieh, G. (2012). The influence of competition, cooperation, and player relationship in a motor performance centered computer game. *Computers in Human Behavior*, 28(6), 2100–2106. doi:10.1016/j.chb.2012.06.014

Plazak, J. and Silver, Z., Inducing Emotion Through Size-Related Timbre Manipulations: A Pilot Investigation.

Poels, K., Hoogen, W. van den, Ijsselstein, W., and de Kort, Y. (2012). Pleasure to Play, Arousal to Stay: The Effect of Player Emotions on Digital Game Preferences and Playing Time. *Cyberpsychology, Behavior, and Social Networking*, 15(1), 1–6. Retrieved from <http://www.liebertonline.com/doi/abs/10.1089/cyber.2010.0040>

Polar. (2014). H7 Heart Rate Sensor. Available from: [http://www.polar.com/au-en/products/accessories/H7\\_heart\\_rate\\_sensor](http://www.polar.com/au-en/products/accessories/H7_heart_rate_sensor). Last accessed: 12th August 2014.

Posner, J., Russell, J. A. and Peterson, B. S. (2005) 'The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology.', *Development and psychopathology*, 17(3), pp. 715–34. doi: 10.1017/S0954579405050340.

Potts, G. F., Martin, L. E., Kamp, S.-M. and Donchin, E. (2011) 'Neural response to action and reward prediction errors: Comparing the error-related negativity to behavioral errors and the feedback-related negativity to reward prediction violations', *Psychophysiology*, 48(2), pp. 218–228. Available at: <http://doi.wiley.com/10.1111/j.1469-8986.2010.01049.x>.

Proctor, J.D., 1998. The social construction of nature: Relativist accusations, pragmatist and critical realist responses. *Annals of the Association of American Geographers*, 88(3), pp.352-376.

Przybylski, A. K., Deci, E. L., Rigby, C. S. and Ryan, R. M. (2014) 'Competence-impeding electronic games and players' aggressive feelings, thoughts, and behaviors.', *Journal of Personality and Social Psychology*, 106(3), pp. 441–457. Available at: <http://doi.apa.org/getdoi.cfm?doi=10.1037/a0034820>.

Przybylski, A. K., Rigby, C. S. and Ryan, R. M. (2010) 'A motivational model of video game engagement.', *Review of General Psychology*, 14(2), pp. 154–166. Available at: <http://doi.apa.org/getdoi.cfm?doi=10.1037/a0019440>.

Ramshur, J.T., 2010. Design, evaluation, and application of heart rate variability software (HRVAS). Master's thesis. The University of Memphis, Memphis, TN: Master's thesis.

Ramshur, J.T., 2010. HRVAS: Heart Rate Variability Analysis Software. University of Memphis. Department of Biomedical Engineering.

Ravaja, N. and Kivikangas, J. M. (2008) 'Psychophysiology of digital game playing : The relationship of self- reported emotions with phasic physiological responses', in *Proceedings of Measuring Behavior*. Maastricht, The Netherlands.

Ravaja, N., 2009. The psychophysiology of digital gaming: The effect of a non co-located opponent. *Media Psychology*, 12(3), pp.268-294.

Ravaja, N., Lab, M. I. N. D., Ckir, /, Salminen, M., Holopainen, J., Saari, T., Laarni, J. and Järvinen, A. (2004) 'Emotional Response Patterns and Sense of Presence during Video Games: Potential Criterion Variables for Game Design', pp. 339–347.

Ravaja, N., Saari, T., Salminen, M., *et al.* (2007). Phasic emotional reactions to video game events: a psychophysiological investigation. *Media Psychology*, 8, 343–367.

Ravaja, N., Saari, T., Turpeinen, M., Laarni, J., Salminen, M. and Kivikangas, M. (2005). Spatial presence and emotions during video game playing: does it matter with whom you play? *Presence: Teleoperators and Virtual Environments*, 15(4), 381-392. Doi: 10.1162/pres.15.4.381

Ravaja, N., Saari, T., Turpeinen, M., Laarni, J., Salminen, M. and Kivikangas, M., 2006. Spatial presence and emotions during video game playing: Does it matter with whom you play?. *Presence: Teleoperators and Virtual Environments*, 15(4), pp.381-392.

Rigby, S. and Ryan, R., 2007. The Player Experience of Need Satisfaction (PENS): An applied model and methodology for understanding key components of the player experience. Retrieved from [immersyve.com/PENS\\_Sept07.pdf](http://immersyve.com/PENS_Sept07.pdf).

Rigby, S. and Ryan, R.M., 2011. Glued to games: How video games draw us in and hold us spellbound: How video games draw us in and hold us spellbound. ABC-CLIO.

Robinson, M. D. and Clore, G. L. (2002) 'Belief and feeling: evidence for an accessibility model of emotional self-report.', *Psychological bulletin*, 128(6), pp. 934–960. doi: 10.1037/0033-2909.128.6.934.

Russell, J. A. and Mehrabian, A. Evidence for a three- factor theory of emotions. *Journal of Research in Personality*, 11, (1977), 273-294.

Russell, J.A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6), 1161-1178. doi: 10.1037/h0077714

Russell, W. D., and Newton, M. (2008). Short-Term Psychological Effects of Interactive Video Game Technology Exercise on Mood and Attention. *Journal of Educational Technology*. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&andamp>

Russoniello, C.V., O'Brien, K. and Parks, J.M. (2009). EEG, HRV and Psychological Correlates while Playing Bejeweled II: A Randomized Controlled Study. *Studies in Health Technology and Informatics*, 144, 189-192. doi: 10.3233/978-1-60750-017-9-189

Ryan, R. M., Scott Rigby, C., and Przybylski, A. K. (2006). The Motivational Pull of Video Games: A Self-Determination Theory Approach. *Motivation and Emotion*, 30, 347–363. Retrieved from

<http://link.springer.com.ezproxy.utas.edu.au/content/pdf/10.1007/s11031-006-9051-8>

Salminen, M., and Ravaja, N. (2007). Oscillatory brain responses evoked by video game events: the case of super monkey ball 2. *Cyberpsychology and Behavior : The Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society*, 10(3), 330–8. doi:10.1089/cpb.2006.9947

Salter, S., Pittaway, J., Swabey, K., Capstick, M. and Douglas, T. (2012) ‘Tolerance of the ERP Signatures of Unfamiliar versus Familiar Face Perception to Spatial Quantization of Facial Images\*’, *Creative Education*, 3, pp. 1–8. Available at: <papers3://publication/uuid/1884459F-160C-4C9A-AB87-29DD3EA8E493>.

Sariyanidi, E., Gunes, H. and Cavallaro, A. (2015) ‘Automatic analysis of facial affect: A survey of registration, representation, and recognition’, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(6), pp. 1113–1133. doi: 10.1109/TPAMI.2014.2366127.

Schell, J. (2014) *The Art of Game Design*, Second Edition. Available at: <http://www.barnesandnoble.com/w/the-art-of-game-design-second-edition-jesse-schell/1118176022?ean=9781466598645>.

Scherer, K. R. (2005) ‘What are emotion? And how can they be measured?’, *Social Science Information Sur Les Sciences Sociales*, 44(4), pp. 695–729. doi: 10.1177/0539018405058216.

Seo, S. H., Lee, C. K., and Yoo, S. K. (2013). Neurophysiological and Behavior Responses in Emotion Regulation. *International Journal of Bio-Science and Bio-*

Technology, 5(5), 1–12. Retrieved from

[http://www.sersc.org/journals/IJBSBT/vol5\\_no5/1.pdf](http://www.sersc.org/journals/IJBSBT/vol5_no5/1.pdf)

Sim, T., Gentile, D. A., Bricolo, F., Serpelloni, G., and Gulamoydeen, F. (2012). A Conceptual Review of Research on the Pathological Use of Computers, Video Games, and the Internet. *International Journal of Mental Health and Addiction*, 10(5), 748–769. Retrieved from <http://link.springer.com/10.1007/s11469-011-9369-7>

Sims, A. (1988). *Symptoms in the mind: An introduction to descriptive psychopathology*. London, UK: Bailliere Tindall.

Skosnik, P.D., Chatterton, R.T., Swisher, T. and Park, S., 2000. Modulation of attentional inhibition by norepinephrine and cortisol after psychological stress. *International Journal of Psychophysiology*, 36(1), pp.59-68.

Smith, T.W. and Pittman, T.S., 1978. Reward, distraction, and the overjustification effect. *Journal of Personality and Social Psychology*, 36(5), p.565.

STAT Medical Education. (2014). EMS STAT 12-LEAD. Available: <http://www.statmedicaleducation.com/11-2/ems-stat-12-lead/>. Last accessed: 12th November 2014.

Sukel, K. (2011). *Neuroanatomy - A Primer*. The DANA Foundation. Available: <http://www.dana.org/News/Details.aspx?id=43515>. Last accessed: 25th July 2014.

Suresh, K. P. (2011). An overview of randomization techniques: An unbiased assessment of outcome in clinical research. *Journal of Human Reproductive Sciences*, 4(1), 8-11.



Sweetser, P. and Wyeth, P. (2005) 'GameFlow: A Model for Evaluating Player Enjoyment in Games', pp. 1–24. Available at: [papers3://publication/uuid/F00C7AE4-B5A4-489C-B821-A7D72644495D](https://papers3://publication/uuid/F00C7AE4-B5A4-489C-B821-A7D72644495D).

Tadel, F., Baillet, S., Mosher, J. C., Pantazis, D. and Leahy, R. M. (2011) 'Brainstorm: A user-friendly application for MEG/EEG analysis', *Computational Intelligence and Neuroscience*, 2011(May 2017). doi: 10.1155/2011/879716.

Tashakkori, A. and Teddlie, C. eds., 2010. Sage handbook of mixed methods in social and behavioral research. Sage.

Tauer, J. M. and Harackiewicz, J. M. (2004) 'The Effects of Cooperation and Competition on Intrinsic Motivation and Performance.', *Journal of Personality and Social Psychology*, 86(6), pp. 849–861. Available at: <http://doi.apa.org/getdoi.cfm?doi=10.1037/0022-3514.86.6.849>.

Tetris. (2010). Tetris Party Deluxe. Nintendo Wii video game. Available at: [http://www.tetrisparty.com/deluxe/en\\_us/](http://www.tetrisparty.com/deluxe/en_us/). Last accessed: 17th November 2014.

Texas Tech. (2014). Psychophysiology Lab, Texas Tech University. Available from: <http://www.depts.ttu.edu/comc/graduate/psychophysiology.php>. Last accessed 27th November 2014.

Trans Cranial Technologies (2012). 10/20 System Positioning. (pp. 1–19). Wanchai, Hong Kong.

Tuli, F., 2011. The basis of distinction between qualitative and quantitative research in social science: reflection on ontological, epistemological and methodological perspectives. *Ethiopian Journal of Education and Sciences*, 6(1).

Vachiratamporn, V., Legaspi, R., Moriyama, K. and Numao, M., 2013, September. Towards the design of affective survival horror games: An investigation on player affect. In *Affective Computing and Intelligent Interaction (ACII)*, 2013 Humaine Association Conference on (pp. 576-581). IEEE.

van den Hoogen, W., Poels, K., IJsselsteijn, W. a. and de Kort, Y. a W. (2012) 'Between Challenge and Defeat: Repeated Player-Death and Game Enjoyment', *Media Psychology*, 15(4), pp. 443–459. doi: 10.1080/15213269.2012.723117.

Vanden Abeele, V., Nacke, L.E., Mekler, E.D. and Johnson, D., 2016, October. Design and preliminary validation of the player experience inventory. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts* (pp. 335-341). ACM.

Vansteenkiste, M. and Deci, E. L. (2003) 'Competitively contingent rewards and intrinsic motivation: Can losers remain motivated?', *Motivation and Emotion*, 27(4), pp. 273–299. Available at: <http://link.springer.com/10.1023/A:1026259005264>.

Vorderer, P., Hartmann, T. and Klimmt, C. (2003) 'Explaining the enjoyment of playing video games: the role of competition', in *ICEC '03: Proceedings of the second international conference on Entertainment computing*. Carnegie Mellon University. Available at:  
<http://portal.acm.org/citation.cfm?id=958720.958735andcoll=DLanddl=GUIDEandCFID=339422392andCFTOKEN=46619683>.

Vorderer, P., Klimmt, C. and Ritterfeld, U., 2004. Enjoyment: At the heart of media entertainment. *Communication theory*, 14(4), pp.388-408.

Walsham, G., 2006. Doing interpretive research. *European journal of information systems*, 15(3), pp.320-330.

Wang, X., and Perry, A. C. (2006). Metabolic and physiologic responses to video game play in 7- to 10-year-old boys. *Archives of Pediatric and Adolescent Medicine*, 160(4), 411-415.

Watling, C.J. and Lingard, L., 2012. Grounded theory in medical education research: AMEE Guide No. 70. *Medical teacher*, 34(10), pp.850-861.

Watson, D., Clark, L. A., and Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social ...* Retrieved from <http://psycnet.apa.org/psycinfo/1988-31508-001>

Weber, M., 2009. *The theory of social and economic organization*. Simon and Schuster.

Weber, R., Ritterfeld, U., and Mathiak, K. (2006). Does Playing Violent Video Games Induce Aggression? Empirical Evidence of a Functional Magnetic Resonance Imaging Study. *Media Psychology*, 8(1), 39–60. doi:10.1207/S1532785XMEP0801\_4

Wehbe, R.R. and Nacke, L., 2013. An Introduction to EEG Analysis Techniques and Brain-Computer Interfaces for Games User Researchers. In *DiGRA Conference*.

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, 92, 548-573.

Weinfurt, K.P., 2000. Repeated measures analysis: ANOVA, MANOVA, and HLM.

Wu, S., Cheng, C. K., Feng, J., Angelo, L. D., Alain, C., and Spence, I. (2012). Playing a First-person Shooter Video Game Induces Neuroplastic Change. *Journal of Cognitive Neuroscience*, 24(6), 1286–1293.

YCAN Australia. (2014). NeXus Skin Conductance Sensor. Available: <http://www.ycanaustralia.com/SKIN-CONDUCTANCE-SENSOR-FOR-NEXUS-4-10-32F>. Last accessed: 17th November 2014.

Yee N., Ducheneaut N., Nelson L. (2012). Online gaming motivations scale: development and validation, in Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, (Austin, TX). Available online at: <http://nickyee.com/pubs/2012%20CHI%20-%20Motivations%20Scale.pdf>

Zammitto, V., Mirza-Babaei, P., Livingston, I., Kobayashi, M. and Nacke, L. E. (2014) 'Player Experience: Mixed Methods and Reporting Results', CHI '14 Extended Abstracts on Human Factors in Computing Systems, pp. 147–150. doi: 10.1145/2559206.2559239.

## Appendices

### Appendix A: Alternative Video Games Considered

Prior to the selection of Tetris Party Deluxe, a number of other video games were considered for use within this research project. There were too many games considered to discuss all of them, so only the top contenders will be briefly outlined below, including rationale for their rejection.

#### 1: Super Monkey Ball 2

Super Monkey Ball 2 (Sega, 2002) is a popular multiplayer video game available on the Nintendo GameCube gaming platform that was considered due to its use in a number of publications investigating social video game play (Ravaja *et al.*, 2004; Salminen and Ravaja, 2007), because of its high Metacritic rating of 87 (Metacritic, 2015), along with its use in peripheral research such as being used as a training tool for surgeons (Rosser *et al.*, 2007). The multiplayer mode involves players competing to be first to complete a series of levels, with each level being comprised of a series

of obstacles, platforms, and challenges that require dexterity and skill to perform without falling off.



Figure 100: A racing level in Super Monkey Ball 2, showing a series of platforms and obstacles (Sega, 2003).

Super Monkey Ball 2 was rejected due to it not being available on modern gaming platforms, which would have made it difficult to integrate within the testing environment. It also is not so well known as Tetris, and is a faster paced game that requires a higher level of dexterity and manual control compared to other video games considered, which would have made it difficult for many less experienced participants to play.

### 2: Left 4 Dead 2

Left 4 Dead 2 (Valve, 2009) is a popular cooperative first person shooter available on a range of modern gaming platforms that also offers a competitive play mode alongside its multiple cooperative modes. Left 4 Dead 2 was considered due to its

multiple social game modes, along with its high Metacritic rating of 89 (Metacritic, 2009). The core game focuses on four players working together to survive an apocalyptic pandemic and fighting their way through masses of infected non-player characters to proceed through levels and make it to checkpoints.



Figure 101: Four players working together to survive an attacking horde of infected in Left 4 Dead 2 (Valve, 2009).

Left 4 Dead 2 was rejected for a number of reasons, the primary being due to the length of each game run, which could take anywhere between 15 and 45 minutes and would therefore cause experimental sessions to take a large amount of time to complete per participant. Additionally, Left 4 Dead 2 has a feature known as Director 2.0, which varies the gameplay intensity and events dynamically based upon players' performance. While this would potentially have assisted in ensuring that all participants were adequately challenged, it would have made it difficult to assess performance and experiences between participants if the core game experience was different in each case.

### 3: Mario Party 10

Mario Party 10 (Nintendo, 2015) is a popular party video game available on the Nintendo Wii U gaming platform, and involves four players competing through a



series of mini games and challenges. Mario Party 10 was considered as it features a good selection of different game modes and components, is a brand new social video game on a current gaming platform, and has a relatively large player following.



Figure 102: Players competing in a mini-game in Mario Party 10 (Nintendo, 2015).

Mario Party 10 was rejected primarily because of its lower Metacritic rating of 66 (Metacritic, 2015), and because many of the mini-games were considered quite childish or overly simple, and therefore potentially struggle to fully engage adult audiences. These considerations rendered Mario Party 10 inappropriate for the current research project.

## Appendix B: Alternative Subjective Measures of Player Experience

### 1: Player Experience of Needs Satisfaction (PENS)

The Player Experience of Needs Satisfaction (PENS) survey is a statistically reliable measure (Johnson and Gardner, 2010) for five different aspects of the subjective player experience using a series of 7-point Likert scales, with answers ranging between 1: “Do Not Agree” to 7: “Strongly Agree”.

The five aspects examined through the PENS survey are:

- Presence/Immersion: The feeling of immersion within the game, which considers how well the game held players attention, as well as how involved players were emotionally (Rigby and Ryan, 2007).
- Relatedness: How connected or related players feel towards other players within the game (Rigby and Ryan, 2007).
- Autonomy: How in-control the player feels, which includes how free players feel to make choices within the game (Rigby and Ryan, 2007).
- Competence: How proficient the player feels, which includes how capable or skilled they feel within the game (Rigby and Ryan, 2007).
- Intuitive Controls: How intuitive the player feels the game controls were, which considers the ease of use, and the level of frustration of the control interface (Rigby and Ryan, 2007).

The PENS survey may provide insights into the state of flow experienced by players during game play, which is a rewarding mental state where the player is fully engaged, focused, and feeling high levels of enjoyment (Thin, Hansen and McEachen, 2011). The PENS survey may also provide insights into other emotional and motivational states experienced by players during different aspects of game play (Johnson and Gardner, 2010).



While the PENS survey is relatively popular and reliable, it was rejected in this project as it does not cover some of the required areas of the video game play experience, such as positive and negative affect.

### 2: Social Presence in Gaming Questionnaire (SPGQ)

The Social Presence in Gaming Questionnaire is a survey instrument that aims to assess the psychological and behavioural engagement that players have with other players during social gaming interactions, and is based upon the psychological theory of Networked Minds Social Presence (de Kort *et al.*, 2007).

The SPGQ consists of a number of items across three scales, which are assessed on a likert scale ranging from 0 to 4 (de Kort *et al.*, 2007):

- Psychological Involvement (Empathy): The extent to which players' emotions were connected to the emotions of other players, such as feeling happy when other players were happy.
- Psychological Involvement (Negative feeling): The extent to which players' emotions were negatively influenced by social interactions, such as feeling jealous when other players succeed.
- Behavioural Engagement: The extent to which players' felt that their actions were interrelated with the other players, such as feeling that their own actions depended upon the actions of other players.

The SPGQ is a useful and reliable survey instrument, however it was deemed insufficient for the purposes of this research project as it does not sufficiently cover the game experience in isolation from social interactions, i.e. the experience of the game independent from the social aspects. It was therefore considered inappropriate and rejected for the purposes of this study.

### Appendix C: Alternative Subjective Measures of Affect

#### 1: Profile of Mood States (POMS)

The Profile of Mood States (POMS) survey uses up to 65 different adjective statements that are rated through a 5-point Likert scale in order to assess six different aspects of mood and their associated states (Russoniello *et al.* 2009). The POMS survey is able to provide both assessments of both transient or fluctuating feelings, as well as enduring mood states (Heuchert and McNair, 2012), and takes between 5 to 15 minutes depending on the number of items used. Although not commonly used in video game studies, there is potential that this could provide a valuable subjective measure of the relative affective variations elicited through interaction with video games.

The six aspects of mood measured through the POMS are:

- Tension/Anxiety: The state of mental or emotional strain, which can include the feeling of conflict, pressure, or anxiety (Russoniello *et al.* 2009).
- Depression/Dejection: The low mood state of aversion to activity, which can include feeling sad, empty, hopeless, and irritable (Russoniello *et al.* 2009).
- Fatigue/Inertia: The state of feeling mentally drained or tired, which includes feelings of weakness and exhaustion (Russoniello *et al.* 2009).
- Vigor/Activity: The state of being energised or motivated, which includes the feeling of efficacy (Russoniello *et al.* 2009).
- Confusion/Bewilderment: The state of being unclear or indecisive, which includes the feeling of puzzlement (Russoniello *et al.* 2009).

The POMS survey is different from the GEQ, SAM, and PENS surveys, in that it is a general instrument for the measurement of mood states within a wide variety of applications rather than purely the investigation of video games. However, in collaboration with subject matter experts it was decided that the POMS survey would not be appropriate for use in this study as it would require interruptions

during video game play to administer the survey, which contravenes the naturalistic testing objectives of this research project.

With the elimination of the POMS as an option for subjective measurements of affect, and lacking any viable alternatives, it was decided that this project would not utilise a subjective survey instrument to measure affect. This had the additional benefit of drastically reducing the length of time that participants spent completing surveys, thereby reducing potential survey fatigue.

## Appendix D: Fuzzy Logic Rules Used

### 1. Mapping Physiological Signals to AV Space

The rules for mapping physiological signals into the AV space (from Mandryk *et al.*, 2006):

1. If (GSR is high) then (arousal is high);
2. If (GSR is mid-high) then (arousal is mid-high);
3. If (GSR is mid-low) then (arousal is mid-low);
4. If (GSR is low) then (arousal is low);
5. If (HR is low) then (arousal is low);
6. If (HR is high) then (arousal is high);
7. If (GSR is low) and (HR is high) then (arousal is mid- low);
8. If (GSR is high) and (HR is low) then (arousal is mid- high);
9. If (EMGfrown is high) then (valence is very low);
10. If (EMGfrown is mid) then (valence is low);
11. If (EMGsmile is mid) then (valence is high);
12. If (EMGsmile is high) then (valence is very high);
13. If (EMGsmile is low) and (EMGfrown is low) then (valence is neutral);
14. If (EMGsmile is high) and (EMGfrown is low) then (valence is very high);
15. If (EMGsmile is high) and (EMGfrown is mid) then (valence is high);
16. If (EMGsmile is low) and (EMGfrown is high) then (valence is very low);
17. If (EMGsmile is mid) and (EMGfrown is high) then (valence is low);
18. If (EMGsmile is low) and (EMGfrown is low) and (HR is low) then (valence is low);
19. If (EMGsmile is low) and (EMGfrown is low) and (HR is high) then (valence is high);
20. If (GSR is high) and (HR is mid) then (arousal is high);
21. If (GSR is mid-high) and (HR is mid) then (arousal is mid-high);
22. If (GSR is mid-low) and (HR is mid) then (arousal is mid-low).

### 2. Mapping AV Space Data to 5 Classifications of Emotions

The rules for mapping the AV space into the model of emotions (from Mandryk *et al.*, 2006):

1. If (arousal is not veryLow) and (valence is midHigh) then (fun is low);
2. If (arousal is not low) and (valence is midHigh) then (fun is low);
3. If (arousal is not veryLow) and (valence is high) then (fun is medium);
4. If (valence is veryHigh) then (fun is high);
5. If (arousal is midHigh) and (valence is midLow) then (challenge is low);
6. If (arousal is midHigh) and (valence is midHigh) then (challenge is low);
7. If (arousal is high) and (valence is midLow) then (challenge is medium);
8. If (arousal is high) and (valence is midHigh) then (challenge is medium);
9. If (arousal is veryHigh) and (valence is midLow) then (challenge is high);
10. If (arousal is veryHigh) and (valence is midHigh) then (challenge is high);
11. If (arousal is midLow) and (valence is midLow) then (boredom is low);
12. If (arousal is midLow) and (valence is low) then (boredom is medium);
13. If (arousal is low) and (valence is low) then (boredom is medium);
14. If (arousal is low) and (valence is midLow) then (boredom is medium);
15. If (arousal is midLow) and (valence is veryLow) then (boredom is high);
16. If (arousal is low) and (valence is veryLow) then (boredom is high);
17. If (arousal is veryLow) and (valence is veryLow) then (boredom is high);
18. If (arousal is veryLow) and (valence is low) then (boredom is high);
19. If (arousal is veryLow) and (valence is midLow) then (boredom is high);
20. If (arousal is midHigh) and (valence is midLow) then (frustration is low);
21. If (arousal is midHigh) and (valence is low) then (frustration is medium);
22. If (arousal is high) and (valence is low) then (frustration is medium);
23. If (arousal is high) and (valence is midLow) then (frustration is medium);
24. If (arousal is midHigh) and (valence is veryLow) then (frustration is high);
25. If (arousal is high) and (valence is veryLow) then (frustration is high);
26. If (arousal is veryHigh) and (valence is veryLow) then (frustration is high);
27. If (arousal is veryHigh) and (valence is low) then (frustration is high);
28. If (arousal is veryHigh) and (valence is midLow) then (frustration is high);

29. If (valence is veryLow) then (fun is veryLow)(challenge is veryLow);
30. If (valence is low) then (fun is veryLow)(challenge is veryLow);
31. If (valence is high) then (challenge is veryLow)(boredom is veryLow)(frustration is veryLow);
32. If (valence is veryHigh) then (challenge is veryLow) (boredom is veryLow)(frustration is veryLow);
33. If (valence is midHigh) then (boredom is veryLow) (frustration is veryLow);
34. If (arousal is veryLow) then (challenge is veryLow) (frustration is veryLow);
35. If (arousal is low) then (challenge is veryLow)(frustration is veryLow);
36. If (arousal is midLow) then (challenge is veryLow) (frustration is veryLow);
37. If (arousal is midHigh) then (boredom is veryLow);
38. If (arousal is high) then (boredom is veryLow);
39. If (arousal is veryHigh) then (boredom is veryLow);
40. If (arousal is veryLow) and (valence is midHigh) then (fun is veryLow);
41. If (arousal is low) and (valence is midHigh) then (fun is veryLow);
42. If (arousal is veryLow) and (valence is high) then (fun is low);
43. If (valence is midLow) then (fun is veryLow) 66. If (arousal is veryLow) and (valence is high) then (boredom is low);
44. If (arousal is low) and (valence is midHigh) then (boredom is low);
45. If (arousal is veryLow) and (valence is midHigh) then (boredom is medium);
46. If (arousal is veryHigh) and (valence is veryLow) then (challenge is medium);
47. If (arousal is veryHigh) and (valence is veryHigh) then (challenge is medium);
48. If (arousal is high) and (valence is low) then (challenge is low);
49. If (arousal is high) and (valence is high) then (challenge is low);
50. If (arousal is veryHigh) and (valence is low) then (challenge is high);
51. If (arousal is veryHigh) and (valence is high) then (challenge is high);
52. If (arousal is midHigh) and (valence is midHigh) then (excitement is low);
53. If (arousal is high) and (valence is midHigh) then (excitement is medium);
54. If (arousal is high) and (valence is high) then (excitement is medium);
55. If (arousal is midHigh) and (valence is high) then (excitement is medium);
56. If (arousal is veryHigh) and (valence is midHigh) then (excitement is high);

- 57. If (arousal is veryHigh) and (valence is high) then (excitement is high);
- 58. If (arousal is veryHigh) and (valence is veryHigh) then (excitement is high);
- 59. If (arousal is high) and (valence is veryHigh) then (excitement is high);
- 60. If (arousal is midHigh) and (valence is veryHigh) then (excitement is high);
- 61. If (arousal is midLow) then (excitement is veryLow);
- 62. If (arousal is low) then (excitement is veryLow);
- 63. If (arousal is veryLow) then (excitement is veryLow);
- 64. If (valence is veryLow) then (excitement is veryLow);
- 65. If (valence is low) then (excitement is veryLow); and
- 66. If (valence is midLow) then (excitement is veryLow).

### Appendix E: The Process of Registering and Coding Events within the CoE

#### 1. Video Recordings: Facial Expressions

Participant facial expressions were coded using the Facial Expression Coding System (FACES) (Kring and Sloan, 1991), which provides guidelines and tools to assist in the detection and classification of facial expressions based upon the detection, duration, valence, and intensity of observed expressions. Coding of facial expressions inherently requires a degree of subjective interpretation on the part of the coder (Ekman, 1982). However, the use of systems for coding facial expressions can attempt to minimise the influence of the coder and their respective interpretations upon the resultant data (Kring and Sloan, 1991).

The FACES approach to coding facial expressions from videos classifies emotions into generally positive versus negative valences based upon Russell's circumplex model of arousal and valence (Russell, 1989), making it consistent with the methods of processing and analysing other affective data sources within this project.

##### 1.1. Coding Expressions using FACES

Faces uses a Coding Sheet to record details of each facial expression as they occur within recorded videos, as illustrated in Figure 103 below. Each detected expression was coded with respect to its duration, its valence, and its intensity, based upon the guidelines outlined below (Kring and Sloan, 1991). The coded expressions were first recorded within the Coding Sheet, and then subsequently entered into the Catalogue of Events.

**Detection:** Expressions were detected when there were observed movements of the face from neutral display (blank or passive) to a non-neutral display (expression), and then back to neutral display. Where the face did not return immediately to neutral, but instead transitioned into a second expression, this was coded as a secondary display additional to the first. For each expression detected, an entry in the Coding Sheet was created.



Time start _____	Time end _____	Duration _____	
<b>Valence:</b>	Positive ____	Negative ____	
<b>Intensity:</b>	low 1	medium 2	high 3
		very high 4	

*Figure 103: A sample Coding Sheet used within this project, depicting recordings of the time logs of each expression, the calculated duration, estimated intensity of expression, and areas to code expressions as either positively or negatively valenced. Adapted from Kring and Sloan (1991).*

**Duration:** The duration of each expression was recorded by subtracting the end time of each expression from the start time of each expression, and recorded in seconds on the Coding Sheet. From these, the average duration for all negative and all positive expressions were calculated separately by dividing the total durations of expressions by the total number of occurrences coded.

**Valence:** The valence (either positive or negative) was decided by the coder, with an appropriate notation being made on the Coding Sheet. Where the valence was uncertain, a list of affect descriptors of both positively and negatively valenced emotions derived from Russell (1989) was used to assist the coder to make consistent decisions.

- **Positive Affect Descriptors:** Happy, Delighted, Glad, Amused, Pleased, Content, Satisfied, Calm, Serene, Excited, Astonished, Cheerful, Surprised, Active, Content.
- **Negative Affect Descriptors:** Miserable, Distressed, Annoyed, Jittery, Nervous, Angry, Gloomy, Anxious, Afraid, Tense, Alarmed, Frustrated, Disgusted, Depressed, Hostile.

**Intensity:** The intensity of each expression was rated by the coder between 1 and 4 (1 = low, 2 = medium, 3 = high, 4 = very high). Expressions were rated as low if expressions were mild, for example a smile where the participant slightly raises their mouth without showing teeth. Medium intensity expressions required additional facial activation, for example an open-mouthed smile where the eyebrows have also been raised. High intensity expressions were reserved for expressions that activated most of the face, such as an open-mouthed laugh with additional cheek and eyebrow movements. Very high intensity expressions were used for expressions where the participant was extremely expressive, including throwing head back in laughter.

### 2. Video Recordings: Social Interactions

While coding and analysing data against specific types or categories of social interactions would doubtless provide a wealth of insights into the nature of social video game play, it was deemed to add unnecessary complexity to the analysis for this project. Social interactions were therefore simply recorded in terms of their occurrence and duration, in a similar manner to the detection and recording of facial expressions described previously, but no attempts were made to classify interactions in any way.

**Detection:** Social interactions were detected when there were observations of one-way speaking intended for the other participant within the dyad without reply, two way interactions between participants within the dyad, or physical interactions of any kind, such as touching. For each interaction detected, an entry into the Catalogue of Events was created.

**Duration:** The duration of each social interaction was recorded by subtracting the end time of each interaction from the start time of each interaction, and recorded in seconds within the Catalogue of Events. From these, the average duration of all interactions could be calculated for each experimental condition. Additionally, social

interaction events could be used for detailed analysis against physiological responses, facial expressions, and game events.

### 3. Video Recordings: Game Events

The detection and coding of game events facilitated the analysis of affective responses to specific and discrete game events occurring during the video game play time series. Game events of interest were detected and coded based upon a predefined series of guidelines, described below. For each event detected, an entry into the Catalogue of Events was created that included a descriptive name, player ownership of the event, and player performance at the time of the event occurring.

**Detection:** For each participant, game events were detected for any non-movement occurrence in the game. For Tetris Party Deluxe, this meant that typical rotation or movement of Tetriminos were excluded from the Catalogue of Events, as these were deemed too neutral to influence affective responses and would not be of interest for the purposes of this project. Examples of event that were detected include successfully clearing lines of Tetriminos, or erroneous placement of Tetriminos in ways that block off areas of the game board.

**Name:** For each new game event detected, an appropriate name was allocated that was descriptive of the event type. The allocated name was reused for all subsequent events of similar type, wherever appropriate. This facilitated grouping of similar game events.

**Description:** For each game event detected, including those where similar events had been previously recorded, a brief notation was made describing the event. Where relevant, observations were made regarding what happened immediately before or after the event occurring. These descriptions were used for later confirmation of event similarity and reiterator reliability assessment purposes, and as such were not recorded within the Catalogue of Events.

**Ownership:** For each game event detected, ownership was assigned to either the player or the opponent. This allowed game events to be analysed for their respective influences upon affective responses when they occur for the participant being recorded, but also provides insights into the affective responses of the participant when the same events occur for the opponent that the player is competing against. For example, the participant being observed might clear multiple lines of Tetriminos from the game board and have a positive affective response, and so a game event is detected. However, it is plausible that the same participant will experience a different affective response when their opponent clears multiple lines of Tetriminos, and so the same game event is detected but it is noted as having occurred to the opponent of the observed participant. Game events were therefore entered into the Catalogue of Events with a notation for whether it occurred to the observed player or the opponent player.

**Performance:** It is plausible that the current performance of the observed participants against their respective opponents influenced the affective responses elicited during each event. Therefore, for each game event detected, the current performance of the observed participant against their respective opponent was noted. Performance was classified as losing when the observed participant was more than 10% behind their opponent in score, neutral when the observed participant was between 10% behind and 10% ahead their opponent in score, and winning when the observed participant was more than 10% ahead of their opponent in score.

#### 4. Model of Emotions: Affective Responses

In addition to the detection and registration of events based upon facial expressions, social interactions, and game events as described above, the time series generated through the Modelling of Emotions fuzzy logic systems (see Section 3.8) were also visually examined to detect affective responses occurring at any point during game play. This was done to ensure that all affective responses occurring during game play

were recorded within the Catalogue, even those that occurred external to any observed facial expressions, social interactions, or game events.

**Detection:** For each participant, affective responses were detected where any of the modelled outputs from the two fuzzy logic systems varied by more than 10% in apparent response to a stimulus.

For each detected affective response, the time series within the Catalogue of Events was examined to confirm whether the response had already been captured through the previously discussed event generation procedures (facial expression events, social interaction events, and game events).

- Where events had previously been created at this point in time due to detection of another event type (facial expression, social interaction, game event, or some combination therein), the affective response data was added to the relevant fields within the existing event within the Catalogue to avoid the creation of duplicate events.
- Where events had not previously been created at this point in time, the affective response data was added as a new event within the Catalogue. Additional event details were sought through examination of video recordings, to provide stimulus information that may have elicited the event. Where no stimulus was evident, “None Detected” was entered into the fields for facial expressions, social interactions, and game events.

**Fuzzy Logic System Outputs:** To facilitate analysis of affective responses and emotional states for any detected event, the Catalogue of Events contains fields for the mean percentage of the maximum score for each of the outputs from the fuzzy logic systems previously discussed (see Section 3.8.2): Arousal, valence, fun, challenge, boredom, frustration, and excitement.

### *3.9.3.5. Registered Events: Completion and Validation*

Each event now contained within the Catalogue of Events was examined for completion and valid data entry, to ensure that fields within the Catalogue are filled in completely, and done so with the correct data formats.

**Check for Duplicates:** The time series for each entry within the Catalogue were compared to ensure that no two entries existed for the same point in time (and therefore same event). Where duplicate events were discovered, the data contained within their respective fields was merged as appropriate. For example, if one of the duplicate event entries contained data for facial expressions but other fields were empty, and the other duplicate event entry contained data for modelled emotions outputs but other fields were empty, then the data contained within second entry would be copied into the first event entry, and the second entry would be deleted.

**Check for Completion:** Each event contained within the Catalogue was checked to ensure that all fields had relevant data entered. Where empty fields were identified, the data sources for those fields were examined at the relevant participant and point in time, and the fields updated with the relevant data.

- Where the empty fields were for facial expressions, social interactions, and game events, if none were evident at the relevant point in time then the fields were updated to contain “None Detected”. This is due to it being possible for affective responses to occur to stimuli that were not recorded or detected within this project.
- For the modelled emotions output from the fuzzy logic systems, the mean percentage of the maximum values was entered for each of the appropriate fields (arousal, valence, fun, challenge, boredom, frustration, excitement). This is due to modelled emotions outputs being continuously recorded and assessed, and so there were always measures for these available.

**Check for Valid Entry:** The data fields for each event contained within the Catalogue was visually inspected to ensure that all fields had the appropriate data contained within it, and that this data was in the appropriate format. This ensured consistency and appropriateness of the data contained within the Catalogue.